IL CAOS

Le leggi del disordine

A cura di

GIULIO CASATI

LE SCIENZE S.p.A. EDITORE

Il caos

IL CAOS
Le leggi del disordine

A cura di

GIULIO CASATI

LE SCIENZE S.p.A. EDITORE

Redazione
Elena Bernacchi

Il curatore

Giulio Casati è professore ordinario di fisica teorica alla Facoltà di scienze dell'Università degli Studi di Milano e coordinatore del Centro di cultura scientifica «A. Volta» di Como. Autore di oltre 90 pubblicazioni in riviste specializzate internazionali, è membro del Consiglio esecutivo della European Science Foundation e di numerose altre istituzioni scientifiche. È tra l'altro responsabile del programma di cooperazione fra l'Istituto di fisica nucleare di Novosibirsk e il Dipartimento di fisica dell'Università degli Studi di Milano.

In copertina

La fotografia di Paul D. Swanson e Julio M. Ottino mostra la complessa struttura prodotta in un fluido viscoso da una goccia di tracciante fluorescente posta in una regione di mescolamento caotico (si veda l'articolo *Il mescolamento dei fluidi* a pagina 152). Il flusso di mescolamento è stato generato dal moto periodico in versi opposti di due cilindri rotanti eccentrici. Sebbene sia caotico, questo flusso possiede alcune simmetrie. Qui un'«isola» di fluido non mescolato si separa in due isole più piccole nel passare per l'asse di simmetria.

Per le traduzioni

Giuseppe O. Longo: *Come nasce la turbolenza*, *I vetri di spin e lo studio dei mezzi disordinati*;
Simona Panattoni: *Determinismo e predicibilità*, *Determinismo e caos*.

ISBN 88-7004-057-7

Prima edizione: gennaio 1991.

Sommario

Introduzione
di Giulio Casati

Ante mare et terras et, quod tegit omnia, coelum,
Unus erat toto naturae vultus in orbe,
Quem dixere Chaos, rudis indigestaque moles
Nec quicquam nisi pondus iners congestaque eodem
Non bene iunctarum discordia semina rerum

Prima del mare e della terra e del cielo che tutto ricopre,
unico e indistinto era l'aspetto della natura in tutto l'universo,
e lo dissero Caos, mole informe e confusa,
nient'altro che peso inerte,
ammasso di germi discordi di cose mal combinate

Ovidio, *Metamorfosi*, I
(Einaudi, Torino, 1979)

Un volume dedicato al caos in una collana di pubblicazioni scientifiche susciterà in molti lettori sorpresa e meraviglia, ma anche curiosità. In effetti, il caos viene associato ad avvenimenti disordinati, incontrollati, che procedono in maniera confusa senza alcuna regola o logica apparenti: crolli in borsa, rivoluzioni, epidemie; l'opposto di quello che comunemente viene inteso come scienza, dalla quale ci si aspetta la capacità di mettere ordine e di predire, in maniera più o meno accurata, l'evolversi dei fenomeni naturali, ma anche di quelli economici, sociali e così via.

Questo atteggiamento ha radici profonde che risalgono agli albori del pensiero scientifico; l'osservazione del moto ricorrente dei corpi celesti, del cielo stellato e il regolare avvicendarsi delle stagioni hanno spinto l'uomo a cercare le leggi che stavano alla radice di questi fenomeni e a predire fatti, prima considerati misteriosi o terrificanti, come l'apparire delle comete e le eclissi di Sole. Dopo duemila anni di tentativi e di sforzi, l'uomo è arrivato alla scoperta di tali leggi. La grande sintesi operata da Newton implica che la dinamica che regge il moto dei corpi celesti così come quello degli altri corpi che ci circondano è esattamente deterministica: questo significa che passato e futuro sono determinati in maniera univoca dallo stato presente. A poco a poco si è fatta strada la convinzione illuministica che l'uomo potesse essere in grado, almeno in linea di principio, di prevedere e magari controllare lo sviluppo degli eventi futuri. «Datemi lo stato dell'Universo e io vi predirò il futuro» diceva Laplace. Questa fiducia illuministica nella capacità dell'uomo, fondata sul determinismo, di predire sempre meglio l'evoluzione degli eventi, si è rafforzata nei secoli successivi, che sono stati testimoni di entusiasmanti conquiste scientifiche: il trionfo della meccanica, la scoperta dell'elettromagnetismo, l'avvento della relatività e della meccanica quantistica.

Tuttavia, le scoperte di questi ultimi vent'anni, rese possibili anche dai moderni calcolatori elettronici, hanno messo in crisi questa visione ottimistica e hanno minato alle fondamenta una certezza che durava da più di tre secoli. La scoperta principale è che regole o leggi perfettamente deterministiche possono produrre un moto completamente caotico e assolutamente imprevedibile. Questo fenomeno viene indicato come *caos deterministico* (si noti l'apparente contraddizione dei termini) e rappre-

senta la terza grande rivoluzione scientifica di questo secolo, dopo la relatività e la meccanica quantistica.

Ma che cosa è esattamente il caos, come può coesistere con l'ordine imposto dalle leggi, che senso ha, e quale possibilità esiste di studiare eventi che sono completamente casuali, quale può essere l'importanza o l'interesse di tali studi? Questi sono alcuni degli interrogativi ai quali questo volume si propone di dare una risposta. In un certo senso, siamo ritornati in una condizione simile a quella dell'uomo primitivo che percepiva il mondo circostante come caotico, con una differenza: noi oggi possiamo scoprire le leggi del caos. Con un paragone un po' ardito, possiamo affermare che la Natura gioca a dadi, ma questi dadi sono «truccati» e a noi spetta il difficile compito di scoprirne le proprietà.

La considerazione fondamentale è che la possibilità di predire il futuro si fonda sul concetto di determinismo: il futuro è determinato in maniera *univoca* dallo stato presente; per esempio, le leggi della dinamica classica come l'equazione di Newton $F = ma$, sono delle «regole» che determinano univocamente il futuro una volta assegnato lo stato iniziale. Dal punto di vista matematico queste regole sono rappresentate da equazioni differenziali, la cui soluzione, una volta assegnate le condizioni iniziali, determina univocamente il futuro. Queste «regole» ci permettono di calcolare la traiettoria di un missile se conosciamo con precisione velocità e posizione di partenza, il giorno e l'ora della prossima eclissi di Sole e così via. Analogamente la conoscenza delle condizioni atmosferiche di oggi permette, in linea di principio, di determinare le condizioni del tempo tra un anno su una certa regione.

È importante sottolineare che questi concetti non riguardano solo i sistemi fisici; al contrario, si applicano a tutti i fenomeni biologici, chimici, ingegneristici, economici, sociali, la cui evoluzione è descritta da leggi non lineari. La non linearità è una caratteristica tipica dei sistemi complessi ed è alla radice del comportamento caotico dei sistemi deterministici. In effetti, se riflettiamo attentamente, comprendiamo che affermare che il futuro di un sistema è determinato in modo univoco dal suo stato presente non significa che siamo effettivamente in grado di determinarlo. Infatti, le equazioni della dinamica possono non essere «risolubili». Questo traduce in linguaggio matematico il fatto ben noto che il futuro della maggior parte dei sistemi è difficile da predire. Forse la persona «non addetta ai lavori», che osserva con meraviglia i progressi della tecnica moderna, può stupirsi che non si sia capaci di risolvere nemmeno il problema del moto di tre corpi in reciproca interazione. Può apparire quasi scandaloso che, a distanza di tre secoli dalla formulazione delle equazioni di Newton, non sappiamo ancora se il nostro sistema solare è stabile e se la Terra continuerà a girare attorno al Sole oppure se andrà a cadervi sopra o, viceversa, abbandonerà la propria orbita per perdersi negli spazi infiniti. Siamo - è vero - in grado di prevedere in modo approssimato l'orbita di un satellite e di seguirne la posizione istante per istante grazie ai moderni calcolatori elettronici, ma non siamo in grado di ottenere la soluzione esatta della sua equazione del moto e quindi di fare previsioni per tempi arbitrariamente lunghi.

In effetti, una delle maggiori conquiste della dinamica moderna è l'aver compreso che sistemi anche apparentemente molto semplici possono portare a soluzioni talmente complicate da apparire del tutto casuali e pertanto estremamente difficili da prevedere. Vengono quindi a cadere quelle barriere psicologiche dovute a secoli di tradizione che hanno considerato determinismo e caos come concetti contrapposti. Il primo a rilevare questo fatto fu il grande matematico e filosofo Henri Poincaré nei suoi studi di meccanica celeste. Tuttavia solo l'avvento dei moderni calcolatori ha consentito di coglierne la rilevanza e le implicazioni.

Come spesso è accaduto nella storia del pensiero scientifico, una scoperta effettuata in un campo molto astratto e apparentemente lontano dalla realtà ha avuto poi sviluppi e applicazioni pratiche assolutamente impensabili. Ciò vale in modo particolare per lo studio del caos e, più in generale, dei sistemi complessi. Questi studi sono di una natura essenzialmente interdisciplinare e ciò spiega l'estrema velocità con la quale essi hanno interessato tutto lo scibile. Un numero sempre più alto di ricercatori è interessato a essi: i matematici che studiano i sistemi di equazioni differenziali non lineari, gli astronomi che studiano la meccanica celeste, i chimici che analizzano le reazioni unimolecolari, i biologi interessati ai modelli non lineari di reazioni biochimiche o ai modelli di dinamica di popolazione, i fisici impegnati nella meccanica statistica, nella ricerca con gli acceleratori di particelle, nell'interazione plasma-laser per la fusione termonucleare e nella teoria delle particelle elementari; e ancora gli ingegneri aeronautici, elettrotecnici e meccanici interessati ai problemi di oscillazione, vibrazione o turbolenza in sistemi diversi, che vanno dalla progettazione di aerei ad alta prestazione ai sistemi di estrazione del petrolio, gli economisti, i medici, i neurologi e così via. Per esempio la ricerca medica ha portato a scoprire che molti parametri fisiologici variano in modo caotico nell'individuo sano mentre una variazione regolare può essere segno di patologia. È il caso del battito del cuore, che in un individuo sano è solo apparentemente regolare, ma si rivela in realtà molto irregolare a un esame più accurato; al contrario, la mancanza di una componente caotica è tipica di condizioni patologiche. Esempi analoghi includono alcuni tipi di leucemia e disordini maniaco-depressivi. Di recente la teoria del caos

è stata applicata a modelli economici realistici e, motivati da questi studi, gli economisti hanno sviluppato nuovi e potenti test per analizzare serie temporali di dati.

Le considerazioni sopra svolte riguardano l'insorgere del comportamento caotico nel mondo degli oggetti cosiddetti «classici», ovvero degli oggetti e dei fenomeni macroscopici. Secondo la visione attuale però le leggi «vere» della fisica sono quelle della *meccanica quantistica*, mentre la meccanica classica fornisce solo una descrizione approssimata, valida per gli oggetti macroscopici. Occorre pertanto comprendere fino a che punto i risultati ottenuti per le leggi classiche si estendono fino a livello microscopico, nel dominio della fisica atomica e molecolare. In un certo senso, poiché le vere leggi sono quelle della meccanica quantistica, noi dobbiamo ripercorrere da capo il cammino compiuto relativamente alla meccanica classica. Per esempio, anche nella meccanica quantistica si manifesta il caos deterministico?

È probabilmente superfluo sottolineare l'importanza di questo problema: il grande balzo in avanti che la meccanica quantistica ha fatto compiere alla scienza e alla tecnologia è sotto i nostri occhi. È pertanto essenziale, oltre che affascinante, cercare di predire il futuro degli eventi del mondo microscopico.

A questo proposito è interessante osservare che la meccanica quantistica è una teoria *intrinsecamente probabilistica*. Una volta assegnato lo stato di un sistema al tempo t mediante una «funzione di stato» $\psi(t)$ noi siamo in grado di fare affermazioni solamente sulla probabilità che eseguendo una misura su una data grandezza si ottenga un determinato valore. Tuttavia è diverso il discorso relativo alla previsione della evoluzione futura. Infatti anche in meccanica quantistica lo stato $\psi(t)$ del sistema al tempo t è determinato univocamente dallo stato iniziale $\psi(0)$. Problema: siamo in grado, date le leggi del moto e dato $\psi(0)$ di predire $\psi(t)$?

Il fatto straordinario e inaspettato è che, a differenza della meccanica classica, la risposta a questa domanda è positiva: per i sistemi quantistici è possibile, almeno in linea di principio, risolvere le equazioni del moto e predire lo stato futuro $\psi(t)$. Pertanto il quadro che si va delineando è diametralmente opposto a quello che si aveva in precedenza: la meccanica classica è sempre stata considerata come una teoria deterministica; ora abbiamo visto che, a causa dell'insorgere del moto caotico, essa porta a un comportamento statistico. D'altro lato la meccanica quantistica è intrinsecamente probabilistica; tuttavia, grazie al suo carattere di stabilità, risulta essere più predicibile della meccanica classica. Questo affascinante tema costituisce oggi uno degli argomenti di ricerca in più rapido sviluppo.

La domanda che il lettore certamente si pone è quali implicazioni concrete può avere lo studio del caos. In altri termini, a quale utilità può portare il sapere che il comportamento della gran parte dei sistemi deterministici è in realtà così complicato da apparire completamente caotico, e che quindi essi si sottraggono alla nostra capacità di previsione?

Anzitutto abbiamo imparato una lezione molto importante: *leggi semplici non portano necessariamente a comportamenti semplici*. Sarebbe alquanto vantaggioso se questo concetto fosse tenuto presente non solo nelle discipline scientifiche, ma anche nella vita politica ed economica.

Un'altra lezione importante è che variazioni piccole nei parametri di un qualunque sistema non portano, necessariamente, a variazioni piccole nel «risultato», cioè nella evoluzione futura: per esempio, un aumento del 5 per cento nell'inquinamento non sempre porta a un peggioramento solo del 5 per cento nel danno ecologico.

Una delle caratteristiche dello studio dei fenomeni caotici è la enorme potenzialità di unificazione culturale in cui tutta la «filosofia naturale» e le discipline economiche, umanistiche, politiche e sociali sono coinvolte. La natura stessa sembra usare il caos nel suo programma di evoluzione: ogni schema deterministico fallirebbe se utilizzato per la sopravvivenza delle forme di vita in condizioni ambientali in continua trasformazione; la natura, pertanto, genera una quantità enorme di forme di vita attraverso mutazioni casuali e, da questa ampia possibilità di scelta, la selezione naturale trova i candidati che si adattano alle mutate condizioni ambientali.

Lo studio del caos ci ha anche portato a capire che la fiducia illimitata nella capacità umana di predire il futuro non è giustificata: arrivato alle soglie del 2000, l'uomo si è accorto che leggi deterministiche possono dare luogo a una evoluzione caotica. Certamente possiamo chiederci quanto sia in realtà importante l'essere riusciti a prendere coscienza di questa forte limitazione della nostra capacità di predire il futuro. Gli uragani, i crolli in borsa, gli attacchi cardiaci, i terremoti sono eventi al di fuori del nostro controllo e ciò è abbastanza spiacevole. Osserviamo tuttavia che, se da un lato è vero che sovente non siamo in grado di predire l'evoluzione di un determinato fenomeno, dall'altro lo studio del caos ci può dire per quali valori dei parametri, cioè in quali condizioni, possiamo ottenere un tipo di comportamento o il suo opposto. Per esempio, possiamo non essere in grado di predire con esattezza l'evolversi di una popolazione biologica o il diffondersi di una malattia, però possiamo essere in grado di sapere in quali condizioni la popolazione si estingue o la malattia scompare, oppure raggiunge un ciclo stazionario o evolve in modo completamente casuale. La conoscenza di questo fatto è spesso molto più importante della conoscenza esatta della evoluzione futura. Infatti, è sui parametri esterni di un sistema che eventualmente noi possiamo agire, ed è importante sapere, per esempio,

come dobbiamo regolare questi parametri per evitare l'insorgere del caos: più che il numero esatto di pesci che popolano un lago o un fiume ci interessa conoscere il livello al di sotto del quale deve essere mantenuto l'inquinamento per evitare che la popolazione si estingua.

Ma le implicazioni del caos deterministico non sono ancora state studiate a fondo. Vogliamo qui solo sottolineare il fatto che tutte le volte che nel suo faticoso ed esaltante cammino verso la comprensione dei fenomeni naturali, l'uomo si è trovato di fronte a delle limitazioni, ciò è stato l'occasione per nuovi grandi balzi in avanti che hanno comportato un rovesciamento della filosofia precedente. Mi riferisco alla osservazione del valore limite della velocità della luce (che ha portato alla teoria della relatività), alla limitazione sulla precisione delle nostre misure (che ha portato alla meccanica quantistica); la teoria del caos ci ha messo di fronte a una limitazione ancora maggiore: l'impossibilità di prevedere il futuro.

A questo si aggiunge, lo sottolineiamo di nuovo, che questa disciplina coinvolge non solo le scienze naturali, ma la nostra stessa vita di tutti i giorni: dai contorni ancora sfumati comincia a intravvedersi la possibilità di nuovi entusiasmanti sviluppi.

Ci sono più cose in cielo ed in terra, Orazio,
di quante nemmeno sono sognate nella tua filosofia.

Amleto: Atto I, Scena V

Le leggi del caos

Determinismo e predicibilità

di David Ruelle

Le leggi della fisica classica sono deterministiche: se si conosce esattamente lo stato di un sistema in un istante dato, si può determinare lo stato di quel sistema in qualsiasi istante successivo. Questa nozione appare in contraddizione con la nostra esperienza quotidiana, in cui certi fatti sembrano prodursi a caso, in modo imprevedibile. L'opposizione tra caso e determinismo, da lungo tempo oggetto di riflessione da parte di ricercatori e filosofi, resta un argomento di controversia.

A questo conflitto è collegato il problema molto concreto della predicibilità (o prevedibilità), per esempio in meteorologia; in questo caso si tratta di stabilire se è possibile prevedere che tempo farà, ma, più in generale, la prevedibilità è la possibilità di prevedere l'evoluzione di un sistema qualsiasi. In questo articolo analizzeremo alcuni progressi teorici relativi al problema del determinismo, del caso e della prevedibilità. Nel corso del nostro studio prenderemo in considerazione diversi dominî della fisica.

Telescopi e biliardi

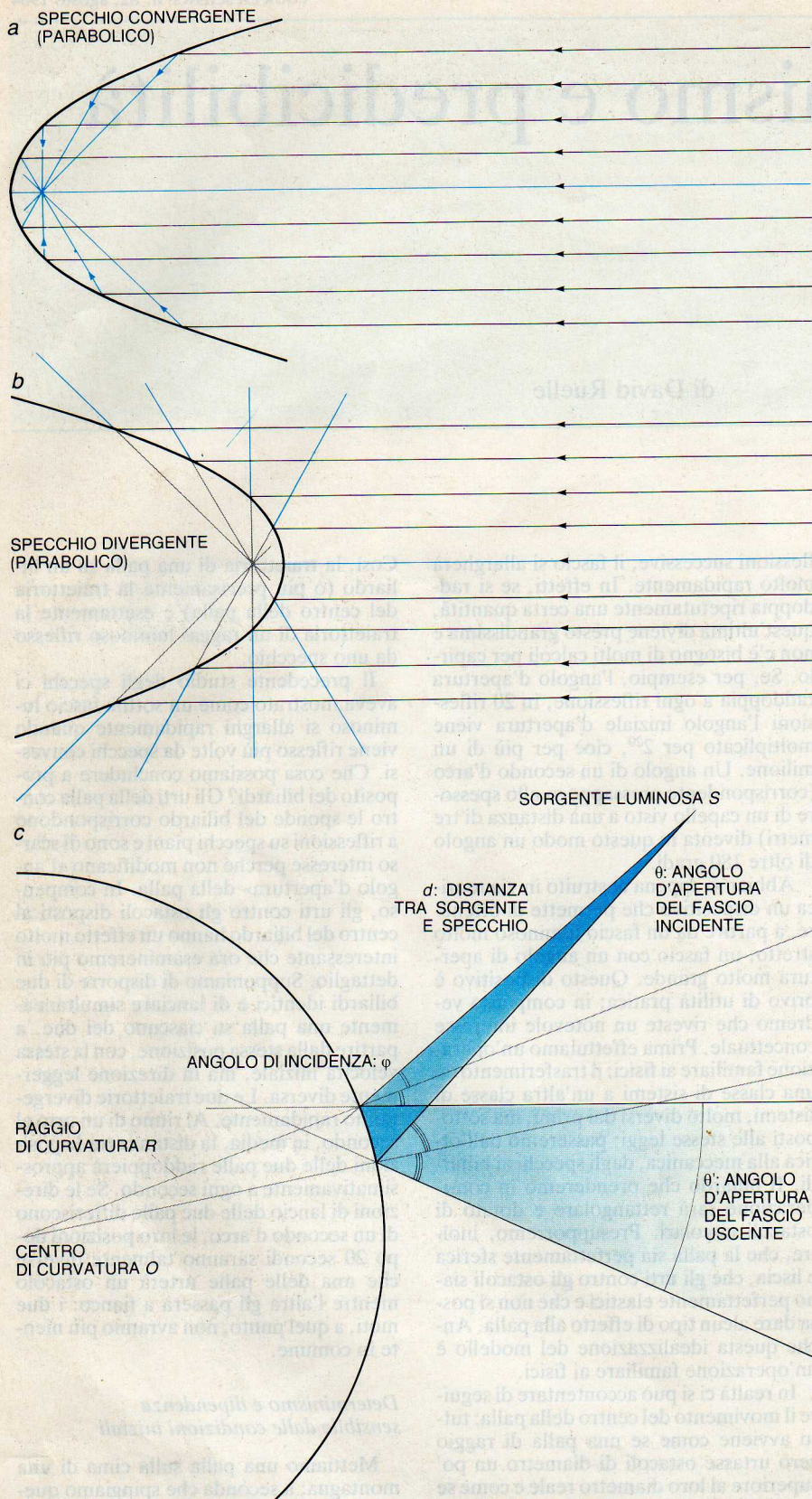
Cominceremo con l'ottica e, più precisamente, con lo studio degli specchi. È noto che nei telescopi uno specchio concavo concentra la luce proveniente dalle stelle su una lastra fotografica. Un fascio di raggi luminosi quasi paralleli viene dunque trasformato in un fascio convergente. Se noi utilizziamo uno specchio convesso, un fascio quasi parallelo diventa invece divergente (*si veda l'illustrazione a pagina 14*). In questa sede ci interesseremo di specchi convessi e, più precisamente, di specchi cilindrici a base circolare. Nell'illustrazione (*c*), un sottile fascio luminoso è riflesso dallo specchio e il fascio emergente diverge più del fascio incidente. In altre parole, l'angolo d'apertura θ' del fascio emergente è maggiore dell'angolo d'apertura θ del fascio incidente. Se ora disponiamo più specchi convessi in modo che il fascio luminoso subisca diverse riflessioni successive, il fascio si allargherà molto rapidamente. In effetti, se si raddoppia ripetutamente una certa quantità, quest'ultima diviene presto grandissima e non c'è bisogno di molti calcoli per capirlo. Se, per esempio, l'angolo d'apertura raddoppia a ogni riflessione, in 20 riflessioni l'angolo iniziale d'apertura viene moltiplicato per 2^{20}, cioè per più di un milione. Un angolo di un secondo d'arco (corrispondente pressappoco allo spessore di un capello visto a una distanza di tre metri) diventa in questo modo un angolo di oltre 180 gradi.

Abbiamo appena costruito in via teorica un dispositivo che permette di ottenere, a partire da un fascio luminoso molto stretto, un fascio con un angolo di apertura molto grande. Questo dispositivo è privo di utilità pratica; in compenso vedremo che riveste un notevole interesse concettuale. Prima effettuiamo un'operazione familiare ai fisici: il trasferimento da una classe di sistemi a un'altra classe di sistemi, molto diversi dai primi, ma sottoposti alle stesse leggi; passeremo dall'ottica alla meccanica, dagli specchi ai biliardi. Il biliardo che prenderemo in considerazione sarà rettangolare e dotato di ostacoli circolari. Presupporremo, inoltre, che la palla sia perfettamente sferica e liscia, che gli urti contro gli ostacoli siano perfettamente elastici e che non si possa dare alcun tipo di effetto alla palla. Anche questa idealizzazione del modello è un'operazione familiare ai fisici.

In realtà ci si può accontentare di seguire il movimento del centro della palla: tutto avviene come se una palla di raggio zero urtasse ostacoli di diametro un po' superiore al loro diametro reale e come se il biliardo fosse un po' meno grande (*si veda l'illustrazione a pagina 15*). (Il centro della palla, infatti, non può avvicinarsi agli ostacoli o alla sponda del biliardo a una distanza inferiore al raggio della palla.) In un urto elastico, l'angolo di incidenza è uguale all'angolo di riflessione, esattamente come avviene per la riflessione di un raggio luminoso su uno specchio. Così, la traiettoria di una palla su un biliardo (o più precisamente la traiettoria del centro della palla) è esattamente la traiettoria di un raggio luminoso riflesso da uno specchio.

Il precedente studio degli specchi ci aveva mostrato come un sottile fascio luminoso si allarghi rapidamente quando viene riflesso più volte da specchi convessi. Che cosa possiamo concludere a proposito dei biliardi? Gli urti della palla contro le sponde del biliardo corrispondono a riflessioni su specchi piani e sono di scarso interesse perché non modificano «l'angolo d'apertura» della palla. In compenso, gli urti contro gli ostacoli disposti al centro del biliardo hanno un effetto molto interessante che ora esamineremo più in dettaglio. Supponiamo di disporre di due biliardi identici e di lanciare simultaneamente una palla su ciascuno dei due, a partire dalla stessa posizione, con la stessa velocità iniziale, ma in direzione leggermente diversa. Le due traiettorie divergeranno rapidamente. Al ritmo di un urto al secondo, in media, la distanza tra le posizioni delle due palle raddoppierà approssimativamente a ogni secondo. Se le direzioni di lancio delle due palle differiscono di un secondo d'arco, le loro posizioni dopo 20 secondi saranno talmente diverse che una delle palle urterà un ostacolo mentre l'altra gli passerà a fianco: i due moti, a quel punto, non avranno più niente in comune.

Determinismo e dipendenza sensibile dalle condizioni iniziali

Mettiamo una palla sulla cima di una montagna; a seconda che spingiamo questa palla a sinistra o a destra, la sua traiettoria sarà ben diversa. Questo è un esempio di «dipendenza sensibile dalle condizioni iniziali»: un piccolo cambiamento nei dati iniziali (spingere la palla a destra o a sinistra) modifica considerevolmente l'evoluzione successiva del sistema. L'esempio della palla sulla montagna potrebbe indurre a pensare che le situazioni ini-

Le modalità di riflessione di un fascio di luce su uno specchio dipendono dalla forma di quest'ultimo. I raggi luminosi riflessi da uno specchio parabolico concavo (*a*) convergono sul fuoco della parabola, mentre su uno specchio parabolico convesso (*b*) il fascio emergente diverge. La divergenza di un fascio di luce avente un piccolo angolo d'apertura aumenta dopo la riflessione su uno specchio sferico convesso (*c*). Più precisamente, chiamando θ l'angolo d'apertura del fascio incidente e θ' l'angolo d'apertura del fascio emergente, si ha la relazione $\theta' = \theta\,(1 + 2d/R\cos\phi)$, nella quale d è la distanza tra la sorgente luminosa S e lo specchio, R è il raggio di curvatura dello specchio e ϕ l'angolo formato dal fascio incidente con la normale allo specchio (angolo d'incidenza).

ziali in cui si manifesta la dipendenza sensibile siano eccezionali, ma l'esempio del biliardo mostra che non è affatto così; per quasi tutte le condizioni iniziali, un piccolo cambiamento di queste condizioni modifica notevolmente l'evoluzione del sistema rispetto a quella di un sistema non perturbato. Inoltre sono numerosi i sistemi che si comportano in questo modo, in pratica tutti i sistemi un po' complessi. Così una piccola modificazione delle condizioni iniziali porta rapidamente a uno stato del tutto differente, il che equivale a dire che una lieve incertezza sui dati iniziali porta rapidamente a una grande incertezza sui risultati. Questo fatto, poco intuitivo, è di notevole importanza filosofica, in quanto mette in luce che un sistema deterministico può avere un comportamento imprevedibile.

Un sistema è deterministico se la conoscenza esatta del suo stato iniziale permette di prevedere il suo futuro con certezza. Questo concetto è stato esposto dall'astronomo e matematico francese Pierre Simon de Laplace (1749-1827) in un celebre testo (*si veda l'illustrazione a pagina 16*). Supponiamo che le leggi della fisica siano deterministiche (torneremo più avanti su questa ipotesi); potremo prevedere nei minimi dettagli il futuro dell'universo, come sostiene Laplace? In pratica, questa affermazione non è né verosimile né vera. Si può ammettere che l'universo, che è un sistema estremamente complesso, presenti il fenomeno della dipendenza sensibile dalle condizioni iniziali. Se la nostra conoscenza dello stato iniziale del sistema è anche di poco incompleta, le nostre previsioni saranno rapidamente soggette a un errore considerevole. Il determinismo, quindi, non implica la predicibilità e il rigore delle leggi fisiche non è in contraddizione con la contingenza dei fatti della vita quotidiana.

Abbiamo detto che l'evoluzione nel tempo di un sistema fisico può rivelare dettagli che la scarsa precisione della conoscenza delle condizioni iniziali non permetteva di prevedere. In certo qual modo, l'evoluzione del sistema ci fornisce informazione, in quanto ci dice quale evoluzione si sia effettivamente realizzata. L'informazione raccolta può essere quantificata; nel caso del biliardo, in cui l'errore raddoppia (in media) a ogni secondo, ciascun secondo trascorso corrisponde alla risoluzione di una scelta binaria: a ogni secondo si produce un bit d'informazione.

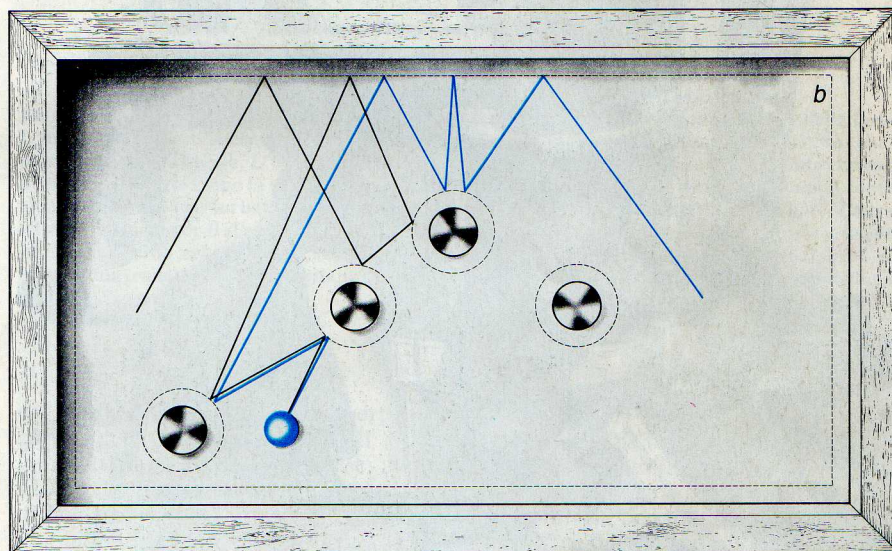
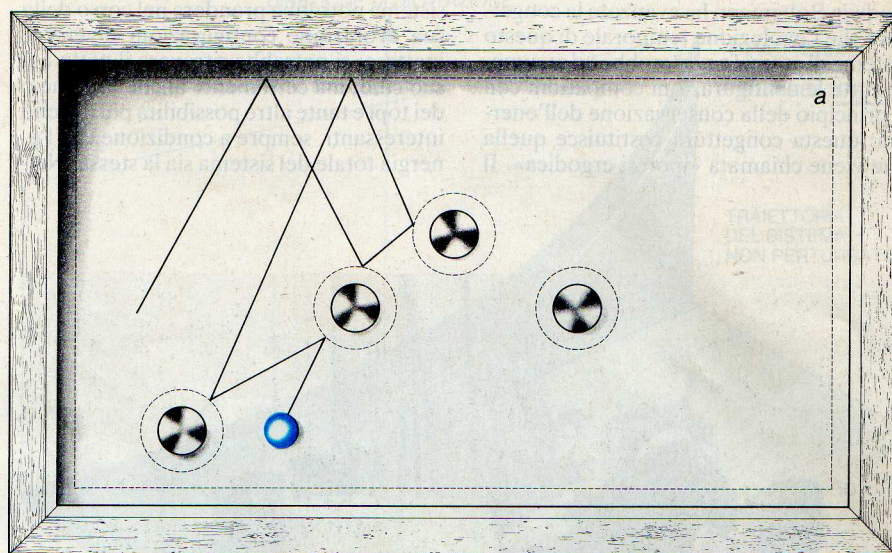
Il lettore ci perdonerà un'osservazione tecnica: abbiamo ammesso che le leggi della fisica siano deterministiche. Questa affermazione, però, va precisata meglio. In effetti, il nostro universo obbedisce più a leggi quantistiche che non alle leggi della meccanica classica che Laplace conosceva. È vero che le leggi quantistiche sono deterministiche, ma esse si applicano a un oggetto matematico, la funzione d'onda, ben diverso dai concetti classici di posizione e velocità di un oggetto materiale. Se si vuole parlare di posizione e di velocità, la funzione d'onda ci permette di ottenere

solo delle probabilità! D'ora in avanti, in questo articolo, dimenticheremo l'aspetto quantistico dell'universo (preponderante nei fenomeni microscopici) e prenderemo in considerazione i problemi più semplici posti dalla meccanica classica. D'altronde, se anche tenessimo conto degli effetti quantistici, le nostre conclusioni non sarebbero molto diverse.

Qualche osservazione di tipo psicologico: è profondamente radicata in noi la coscienza del libero arbitrio, apparentemente incompatibile con un rigoroso determinismo. Inoltre, molti di noi ammetterebbero volentieri la possibilità che forze divine od occulte svolgano un ruolo nel succedersi dei fatti. Così stando le cose, i tentativi di dare un'impronta soggettivistica o indeterministica alla fisica hanno sempre avuto un certo successo. È ciò che è accaduto dopo la scoperta della meccanica quantistica e il problema resta d'attualità. Citiamo, a questo proposito, una recente controversia tra René Thom, Ilya Prigogine e Michel Serres in seguito alla pubblicazione da parte di Thom di un articolo dal titolo provocatorio: *Halte au hasard, silence au bruit*. D'altra parte, il disaccordo tra gli autori appena citati non verte tanto sul fenomeno della dipendenza sensibile dalle condizioni iniziali, quanto piuttosto su considèrazioni filosofiche che non discuteremo in questa sede.

Torniamo ora a un piccolo problema che ha forse creato qualche difficoltà al lettore. Lo stato di un sistema può essere rappresentato da un punto in uno spazio detto «spazio delle fasi». L'insieme degli stati successivi del sistema al variare del tempo forma allora una curva, che è la traiettoria del sistema; l'evoluzione temporale corrisponde allo spostamento lungo la sua traiettoria del punto che rappresenta il sistema. Se si perturba lo stato iniziale del sistema, nel corso dell'evoluzione successiva la traiettoria perturbata si allontana dalla traiettoria non perturbata. Ora, secondo le leggi della meccanica la direzione del tempo non ha un ruolo particolare: si può studiare l'evoluzione verso il passato allo stesso modo dell'evoluzione verso il futuro. Questo fatto è compatibile con le considerazioni precedenti? La situazione è rappresentata nell'illustrazione di pagina 17, in cui si vede come la maggior parte delle piccole perturbazioni di un sistema cresca sia quando il tempo scorre verso il futuro sia quando scorre verso il passato. Ciononostante, certe particolari perturbazioni (nella direzione «di contrazione») diminuiscono verso il futuro, e altre (nella direzione «di dilatazione») diminuiscono verso il passato. Di conseguenza, la dipendenza sensibile dalle condizioni iniziali è compatibile con una perfetta simmetria fra passato e futuro.

L'ipotesi di Boltzmann
e il tempo di Poincaré

Abbiamo appena ricordato la reversibilità del tempo per le leggi della meccanica,

I movimenti di una palla da biliardo ideale corrispondono al percorso di raggi luminosi. In entrambi i casi l'angolo d'incidenza è uguale all'angolo di riflessione: la direzione d'incidenza e la direzione d'uscita formano angoli uguali con la perpendicolare alla sponda urtata dalla palla. Per studiare il moto, si può considerare il percorso del centro della palla su un tavolo da biliardo la cui superficie sia stata ridotta di una striscia larga quanto il raggio della palla (*a*): il centro della palla, infatti, non può avvicinarsi alla sponda a una distanza inferiore al raggio della palla stessa. Analogamente, se si mettono sul biliardo degli ostacoli circolari, si può considerare l'urto del centro della palla contro ostacoli il cui raggio sia aumentato di un valore pari al raggio della palla. Lanciando due palle in direzioni un po' diverse, l'angolo formato dalle due direzioni non sarà modificato dalle riflessioni sulle sponde, ma aumenterà se avvengono urti contro gli ostacoli posti all'interno. Dopo diversi urti di questo genere, i percorsi delle palle potranno evolvere in modo molto differente (*b*).

reversibilità che vale per tutte le leggi fondamentali della fisica. Come mai, allora, il tempo sembra così poco reversibile? Se io lascio andare un bicchiere, questo cade e si rompe, ma non ho mai visto dei pezzi di vetro rotto ricomporsi a formare un bicchiere intatto che salti poi nuovamente nella mia mano. Viviamo in un universo in cui il tempo scorre irreversibilmente, ma i fisici continuano a sostenere che le leggi fondamentali della fisica sono reversibili. Come possono prestare fede a leggi che sembrano violare in tutta evidenza l'osservazione continua e quotidiana di ciascuno di noi?

Per capirlo, prenderemo in considerazione un sistema isolato, per esempio un gatto che mettiamo con del cibo e con tutto ciò che gli può servire in una scatola che lo isoli completamente dal mondo esterno. Il contenuto di questa scatola è composto da un gran numero di elementi microscopici (atomi) soggetti a leggi fisiche, «naturali» e reversibili. Il fisico austriaco

Ludwig Boltzmann ha avanzato la congettura che l'evoluzione temporale di questo sistema di atomi lo porterebbe ad assumere tutte le configurazioni compatibili con il principio della conservazione dell'energia: questa congettura costituisce quella che viene chiamata «ipotesi ergodica». Il sistema potrebbe prendere nel corso della sua evoluzione configurazioni diverse e inattese: si potrebbe ritrovare il gatto e il suo cibo, ma certamente anche un cane o dei topi e tante altre possibilità più o meno interessanti, sempre a condizione che l'energia totale del sistema sia la stessa. Nella maggior parte dei casi, però, la scatola non conterrà che un poco di poltiglia maleodorante.

Seguiamo mentalmente la scatola che all'inizio contiene il gatto. Dopo un certo periodo di tempo, il gatto morirà: lasciamo all'immaginazione del lettore ciò che può accadere nei milioni di anni successivi, ma certamente saremo d'accordo nel ritenere che il gatto non resusciterà. Secondo Boltzmann, invece, si può ritrovare nella scatola un gatto vivo, a condizione di attendere per un tempo abbastanza lungo, quello che oggi viene chiamato tempo di ricorrenza di Poincaré. L'intuizione ci dice che questo periodo è considerevolmente lungo e le stime che se ne possono fare confermano questa idea.

Vediamo ora come si possa conciliare un fenomeno irreversibile - la morte del gatto - con la reversibilità delle leggi naturali della fisica. Partendo dallo stato «gatto», stato speciale e molto improbabile eppure reale, la scatola evolve verso altri stati numerosissimi e tutti differenti nella loro disposizione microscopica, la maggior parte dei quali apparirà semplicemente come un po' di poltiglia. Se si aspetta abbastanza a lungo, la materia riformerà il gatto, ma il tempo necessario è tale che non abbiamo alcuna possibilità d'osservare questo «miracolo».

È possibile fornire dell'irreversibilità una spiegazione più precisa di quella che abbiamo appena tratteggiato: si potrebbero definire i concetti di entropia, di disordine ecc. Questa spiegazione era accettata da Albert Einstein e lo è ancora dalla maggior parte dei fisici. Va notato, però, che recentemente è stata riconsiderata da Ilya Prigogine dell'Università di Bruxelles. In termini matematici, la congettura di Boltzmann significa che l'evoluzione temporale della scatola fa passare il suo punto rappresentativo vicino a tutte le configurazioni possibili; ma non tutte saranno raggiunte. L'evoluzione, dunque, consiste in un eterno quasi-ritorno anziché in un eterno ritorno. Il matematico Yasha Sinaï, di Mosca, ha dimostrato l'ergodicità (cioè la validità dell'ipotesi ergodica) del biliardo a ostacoli convessi che abbiamo illustrato in precedenza.

Questa dimostrazione è una notevole impresa matematica che ci permette di utilizzare la nozione di dipendenza sensibile dalle condizioni iniziali che si realizza in questo sistema. Sinaï ha anche avviato la dimostrazione dell'ergodicità di un sistema costituito da sfere soggette, nel corso dei loro spostamenti, a collisioni perfettamente elastiche. Questo sistema è il modello di un gas monoatomico. La dimostrazione dell'ergodicità di un simile gas «di sfere dure» non è terminata e ci si è inoltre accorti che, contrariamente a quanto pensava Boltzmann (e un gran numero di fisici e matematici dopo di lui), l'ergodicità non è affatto un fenomeno universale: i matematici Anton Kolmogorov e Vladimir Arnold, a Mosca, e Jürgen Moser, attualmente a Zurigo, hanno dimostrato che numerosi sistemi non sono

Essai philosophique sur les probabilités
di Pierre Simon de Laplace

Un'intelligenza che, per un istante dato, potesse conoscere tutte le forze da cui la natura è animata, e la situazione rispettiva degli esseri che la compongono, e che inoltre fosse abbastanza grande da sottomettere questi dati all'analisi, abbraccerebbe nella stessa formula i movimenti dei più grandi corpi dell'universo e quelli dell'atomo più leggero: nulla le risulterebbe incerto, l'avvenire come il passato sarebbe presente ai suoi occhi. Lo spirito umano offre, nella perfezione che ha saputo dare all'astronomia, una debole parvenza di questa intelligenza.

Pierre Simon de Laplace (1749-1827) concepiva il determinismo come la possibilità di conoscere esattamente lo stato futuro di un sistema a partire dalla determinazione precisa del suo stato attuale. In realtà, il fenomeno della dipendenza sensibile dalle condizioni iniziali impedisce di prevedere lo stato futuro dei sistemi, contrariamente a ciò che pensava Laplace. In effetti, due sistemi che inizialmente presentano solo leggere differenze finiscono ben presto per non avere più niente in comune; pertanto i concetti di determinismo e di predicibilità devono essere nettamente distinti.

ergodici. Questo fatto non invalida la spiegazione dell'irreversibilità che abbiamo dato in precedenza, in quanto tale spiegazione dipende fondamentalmente dalla considerevole estensione del tempo di ricorrenza di Poincaré.

Non descriveremo i difficili lavori di Kolmogorov, Arnold e Moser, ma prenderemo in considerazione i risultati ottenuti su un esempio semplice e pittoresco di sistema non ergodico. Su un vassoio quadrato di 182 millimetri di lato abbiamo messo 81 monete da un cruzeiro brasiliano del 1980 (queste monete sono dischi di acciaio con un diametro di 20 millimetri). Il vassoio e le monete costituiscono un modello accettabile di un gas di «sfere dure» a due dimensioni (*si veda* a *nell'illustrazione di pagina 18*). È consentito far scivolare le monete sul vassoio a condizione che non si sovrappongano e non escano dal vassoio. Per prima cosa si possono raggruppare bene le monete per controllare che esse non riempiano completamente il contenitore (*si veda* b *nell'illustrazione di pagina 18*). Questo sistema è ergodico? No, e adesso vediamo perché. Formiamo, con le stesse monete, una griglia quadrata (*si veda* c *nell'illustrazione di pagina 18*): esse possono ancora muoversi un poco, ma nessuno spostamento continuo sul piano permette di passare dalla configurazione *a* alla configurazione *c*. In questa configurazione, nessuna moneta può spostarsi di più di qualche millimetro dalla sua posizione media. L'evoluzione temporale di un gas di sfere dure, quindi, non permette di passare dalla configurazione *c* alla configurazione *a*: non c'è ergodicità.

Il lettore esperto di matematica cercherà di dimostrare questa affermazione considerando un vassoio quadrato di lato pari a (180 + ε) millimetri, con ε abbastanza piccolo. Il lettore amante del fai-da-te raccoglierà le necessarie monete circolari a bordo liscio, identiche l'una all'altra, e costruirà un vassoio quadrato di dimensione adeguata e con un piccolo bordo rialzato. Notiamo inoltre che anche la configurazione *d* è «isolata», cioè irraggiungibile facendo scivolare in modo continuo le monete a partire dalle configurazioni *b* o *c*. Se si ingrandisse un poco il vassoio, lo stretto necessario perché si possa passare dalla configurazione *c* alla configurazione *b*, allora si potrebbe pensare alla configurazione *c* come a un modello della scatola con dentro il gatto di cui abbiamo parlato in precedenza (per i fisici, si tratta di uno stato metastabile). Se il gatto muore, ci vorrà molto tempo perché i moti disordinati lo facciano resuscitare.

Lo stato di un sistema può essere rappresentato da un punto nel cosiddetto «spazio delle fasi». L'insieme degli stati successivi di un sistema forma allora, al variare del tempo, una curva, che è la traiettoria del sistema. L'evoluzione temporale corrisponde allo spostamento del punto che rappresenta il sistema lungo la sua traiettoria. La traiettoria passante per *A* corrisponde a una condizione iniziale non perturbata, quelle passanti per *B*, *C* e *D* a condizioni iniziali perturbate. La traiettoria passante per *B* si avvicina nel futuro alla traiettoria non perturbata. La traiettoria passante per *D* si avvicina nel passato alla traiettoria non perturbata, mentre la perturbazione tipica *C* dà luogo a una traiettoria che si allontana sia nel futuro sia nel passato dalla traiettoria non perturbata.

Le radici del caso

I biliardi e i sistemi di sfere dure che abbiamo studiato non sono i primi sistemi in cui sia stata individuata la dipendenza sensibile dalle condizioni iniziali. In effetti, questa proprietà era stata dimostrata alla fine del secolo scorso dal matematico francese Jacques Hadamard in un modello un po' difficile da visualizzare: quello del «flusso geodetico su una superficie a curvatura costante negativa». Il fisico Pierre Duhem, anch'egli francese, aveva rilevato la portata filosofica del risultato di Hadamard e nel suo libro *La théorie physique. Son objet et sa structure* (Éditions Chevalier et Rivière, 1906) c'è un paragrafo intitolato «Esempio di deduzione matematica per sempre inutilizzabile», in cui Duhem mostra come la dipendenza sensibile dalle condizioni iniziali renda illusorie le previsioni a lungo termine per i sistemi del tipo considerato da Hadamard. Anche il grande matematico francese Henri Poincaré si occupò del problema, come si vede nel suo libro *Science et méthode* (Éditions Ernest Flammarion, Parigi 1908). Poincaré introduce già il gas di sfere dure e la predicibilità dei fenomeni atmosferici (*si veda l'illustrazione a pagina 19*). Per numerosi sistemi, dunque, non siamo in grado di fare previsioni precise a lungo termine, e questo giustifica la nozione empirica di caso. È un punto, questo, su cui Poincaré insiste in modo particolare.

Il caso s'insinua nei sistemi deterministici in quanto un cambiamento impercettibile nelle condizioni iniziali porta, dopo qualche tempo, a cambiamenti notevoli. Lo abbiamo visto per il biliardo. Lo stesso fenomeno appare sotto una luce un po' diversa nei sistemi con molteplici bacini d'attrazione separati da una frontiera complessa. Di che cosa si tratta? Abbiamo già illustrato la situazione in cui una palla posta in cima a una montagna va a finire in valli ben lontane l'una dall'altra se spinta a sinistra oppure a destra. In modo più realistico, possiamo immaginare due bacini fluviali separati da una catena di montagne molto frastagliate. Se la palla cade da un aereo sulla cresta montuosa, sarà difficile prevedere in quale bacino andrà a finire.

Dopo questo esempio semplice ma un po' artificiale, ecco un esempio più soddisfacente. Un pendolo magnetico è una piccola calamita appesa a una barra rigida e libera di muoversi intorno a un punto fisso con un leggero attrito. Sotto questo pendolo mettiamo molte altre piccole calamite che attraggono o respingono quella appesa. In generale, esistono diverse posizioni di equilibrio. A quale di queste posizioni arriverà il sistema se lo si lascia evolvere da una posizione non di equilibrio? Se si realizza in pratica l'esperimento, si vede il pendolo oscillare irregolarmente facendo talvolta brusche deviazioni. A volte sembra fermarsi, poi riparte fino a immobilizzarsi definitivamente in una delle posizioni di equilibrio, che è inizialmente imprevedibile: le diverse posizioni di equilibrio hanno «bacini di attrazione» con frontiere comuni molto com-

plesse, come una catena montuosa molto frastagliata che separa diversi bacini fluviali. Se non si conoscono con estrema precisione la posizione e la velocità iniziale del pendolo, lo stato d'equilibrio finale non è in pratica prevedibile.

L'imprevedibilità di un lancio di dadi è indubbiamente dello stesso tipo. Per poterne dire di più, bisognerebbe disporre di un modello del bussolotto e precisare il modo in cui lo si agita prima di lanciare i dadi. Per quanto riguarda la roulette, non vi sono certamente né dipendenza sensibile dalle condizioni iniziali né bacini di attrazione separati da frontiere complesse; tuttavia il comando muscolare all'atto del lancio è troppo impreciso perché si possa prevedere la posizione finale della pallina. Come si vede, le radici del caso sono molteplici.

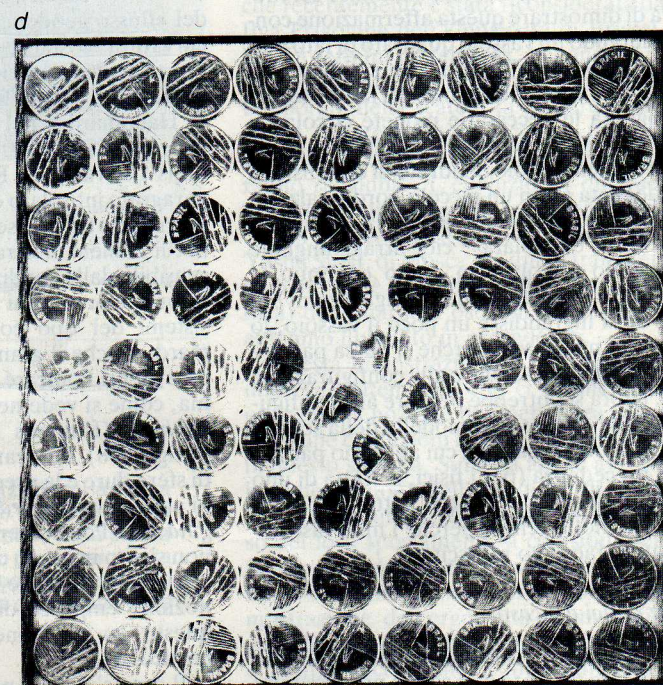
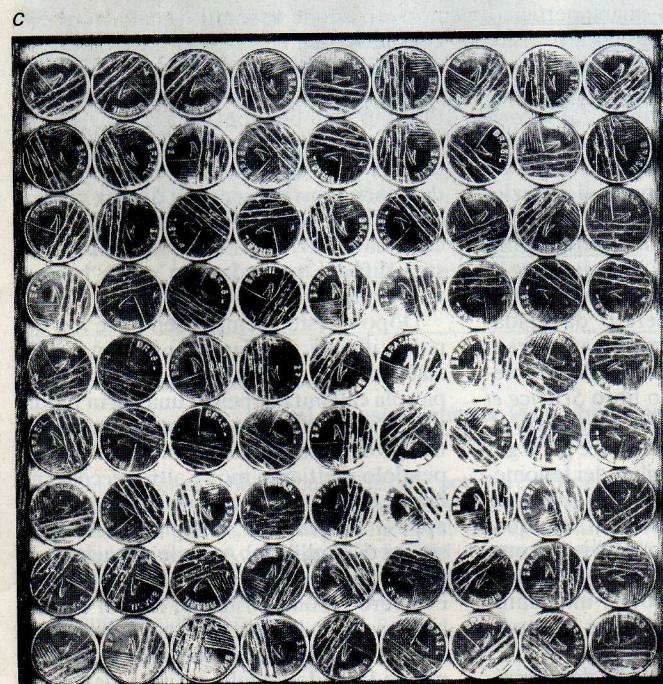
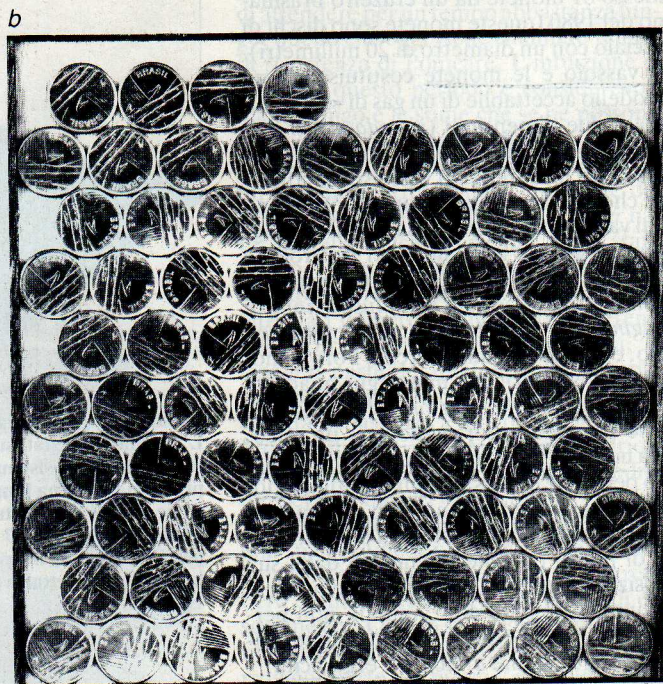
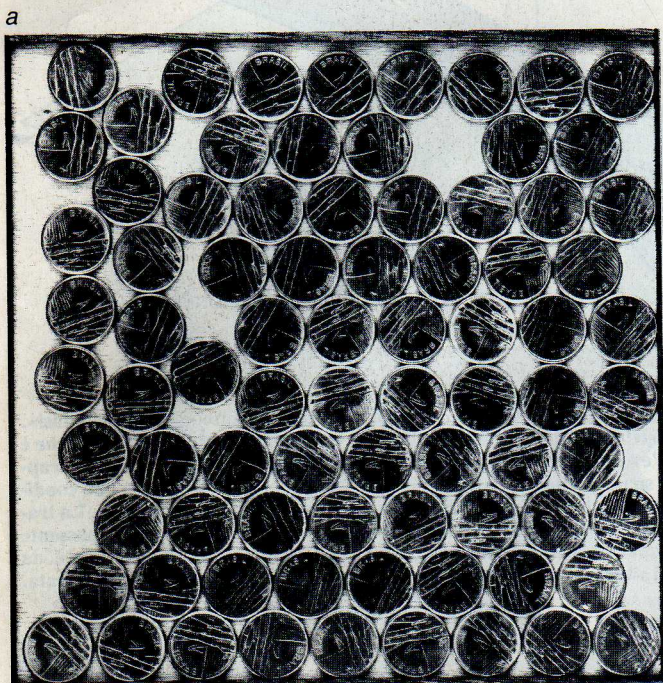
Attrattori strani, turbolenza e predicibilità

Se si agita in modo abbastanza energico dell'aria, dell'acqua o un altro fluido, si producono turbolenze: il movimento del fluido è disordinato e il suo comportamento appare imprevedibile. Così è in effetti, in quanto i sistemi turbolenti presentano il fenomeno della dipendenza sensi-

a

b

c

d

Una serie di monete disposte su un vassoio quadrato costituisce un esempio di sistema non ergodico, vale a dire di sistema che può trovarsi in due stati distinti tali che sia impossibile passare in modo continuo dall'uno all'altro. Ottantun monete di 20 millimetri di diametro sono state sistemate su un vassoio quadrato di 182 millimetri di lato (*a*). Se si avvicinano il più possibile le monete (*b*), si constata che esse non riempiono completamente il vassoio. Se le si dispone in modo che formino una griglia quadrata (*c*), le monete sono ancora in grado di muoversi un poco, ma non esiste alcuno spostamento continuo sul piano che permetta di passare dalla configurazione *a* alla configurazione *c*. Anche la configurazione *d* è «isolata»: infatti è impossibile raggiungerla facendo scivolare in modo continuo le monete a partire dalle configurazioni *b* o *c*.

bile dalle condizioni iniziali. Questo è importante, tra l'altro, perché l'atmosfera terrestre è turbolenta e la predicibilità dei fenomeni meteorologici è fortemente influenzata da questa dipendenza. Henri Poincaré, nel testo riportato nell'illustrazione in questa pagina, sembra averlo capito bene; eppure, le sue idee di tipo troppo qualitativo furono dimenticate e non trovarono posto nelle teorie sulla turbolenza in auge ancora fino a tempi recenti. La teoria che riscuoteva il maggiore favore una decina di anni fa, quella del sovietico Lev Landau e del tedesco Eberhard Hopf, era nettamente incompatibile con il concetto di dipendenza sensibile dalle condizioni iniziali. Nel 1963, però, Edward Lorenz, meteorologo del Massachusetts Institute of Technology, aveva pubblicato un modello semplificato di fluido in cui era chiaramente messo in evidenza il fenomeno di rapida crescita degli errori in funzione delle condizioni iniziali. L'ipotesi di Lorenz era che lo stesso fenomeno si producesse anche per la circolazione dei fluidi nell'atmosfera, il che impedirebbe qualsiasi previsione meteorologica a lungo termine.

Nel 1971, l'olandese Floris Takens e io, respingendo la teoria di Landau e Hopf, avanzammo l'ipotesi che la turbolenza idrodinamica sia rappresentata da attrattori strani, oggetti matematici che descrivono evoluzioni temporali con dipendenza sensibile dalle condizioni iniziali (*si veda l'illustrazione a pagina 20*). D'altra parte, queste nuove ipotesi erano verificabili sperimentalmente. Il fatto che non conoscessimo né gli scritti di Lorenz, né le idee di Poincaré, né il dibattito allora in corso a Mosca era caratteristico di quell'epoca. Il nostro articolo ebbe un'accoglienza abbastanza fredda, ma venne confermato nel 1974 dalle simulazioni numeriche di John McLaughlin e Paul Martin della Harvard University e dagli esperimenti idrodinamici di Gunther Ahlers dei Bell Laboratories e di Jerry Gollub e Harry Swinney del City College di New York. In seguito, l'argomento è divenuto di moda e ha dato luogo a numerose pubblicazioni di carattere sia teorico sia sperimentale. Oggi è possibile misurare la velocità di crescita degli errori nei sistemi turbolenti: si può stimare la quantità di informazione prodotta per unità di tempo in tali sistemi e si ha anche un'idea precisa dei meccanismi che danno inizio alla turbolenza.

Da tutti questi lavori, risulta che si può oggi considerare la dipendenza sensibile dalle condizioni iniziali come la definizione di turbolenza idrodinamica. Questa definizione non fornisce una teoria completa della turbolenza, che resta una delle grandi questioni irrisolte della fisica teorica, ma un primo passo è stato compiuto e il problema risulta oggi posto in termini più corretti. Eccoci tornati, si direbbe, a quello che Poincaré già sapeva tre quarti di secolo fa. Questa affermazione è contemporaneamente vera e falsa: sono molti i casi di scoperte moderne che si rivelano previste da grandi uomini del passato. E

Science et méthode
di Henri Poincaré

Una causa piccolissima che sfugga alla nostra attenzione determina un effetto considerevole che non possiamo mancar di vedere, e allora diciamo che l'effetto è dovuto al caso. Se conoscessimo esattamente le leggi della natura e la situazione dell'universo all'istante iniziale, potremmo prevedere esattamente la situazione dello stesso universo in un istante successivo. Ma se pure accadesse che le leggi naturali non avessero più alcun segreto per noi, anche in tal caso potremmo conoscere la situazione iniziale solo approssimativamente. Se questo ci permettesse di prevedere la situazione successiva con la stessa approssimazione, non ci occorrerebbe di più e dovremmo dire che il fenomeno è stato previsto, che è governato da leggi. Ma non sempre è così; può accadere che piccole differenze nelle condizioni iniziali ne producano di grandissime nei fenomeni finali. Un piccolo errore nelle prime produce un errore enorme nei secondi. La previsione diventa impossibile e si ha un fenomeno fortuito.

Come dobbiamo rappresentare un recipiente pieno di gas? Innumerevoli molecole animate da grandi velocità solcano questo recipiente in tutte le direzioni; a ciascun istante, urtano le parete o si scontrano tra loro; e questi urti hanno luogo nelle condizioni più diverse. Ciò che in questo caso ci colpisce non è la piccolezza delle cause, ma soprattutto la loro complessità. Eppure, il primo elemento è ancora presente e ha un ruolo importante. Se una molecola fosse deviata, verso sinistra o verso destra rispetto alla sua traiettoria, di una quantità piccolissima, paragonabile al raggio d'azione delle molecole di un gas, essa eviterebbe una collisione, oppure la subirebbe in condizioni diverse, e questo farebbe variare, magari di 90 o di 180 gradi, la direzione della sua velocità dopo l'urto.

E non è tutto: abbiamo appena visto che è sufficiente deviare la molecola, prima dell'urto, di una quantità infinitamente piccola, perché essa sia deviata, dopo l'urto, di una quantità finita.

Perché i meteorologi hanno tanta difficoltà a prevedere il tempo con un certo grado di esattezza? Perché i rovesci di pioggia, e le tempeste stesse, ci sembrano arrivare a caso, tanto che molte persone trovano del tutto naturale pregare per avere la pioggia o il bel tempo, mentre riterrebbero ridicolo invocare un'eclisse con la preghiera? Noi vediamo che le grandi perturbazioni si producono generalmente nelle regioni in cui l'atmosfera è in equilibrio instabile. I meteorologi sono ben consapevoli che questo equilibrio è instabile, che un ciclone nascerà da qualche parte, ma dove? Non sono in grado di dirlo; un decimo di grado in più o in meno in un punto qualunque e il ciclone scoppia qui e non là, porta le sue devastazioni in contrade che sarebbero state risparmiate. Se si fosse conosciuto questo decimo di grado, si sarebbe potuto prevederlo in anticipo, ma le osservazioni non erano né abbastanza ravvicinate né abbastanza precise, ed è per questo che tutto sembra dovuto all'intervento del caso.

Henri Poincaré (1854-1912) aveva già iniziato a occuparsi, tre quarti di secolo fa, del problema della predicibilità. Egli aveva compreso che il determinismo delle leggi fisiche non permetteva la previsione dei fenomeni naturali. In realtà, l'analisi di Poincaré era puramente qualitativa; al momento attuale i ricercatori sono viceversa in grado di quantificare la predicibilità dei fenomeni.

tuttavia, così come Leonardo da Vinci non avrebbe potuto costruire un aereo, Poincaré non avrebbe potuto elaborare una teoria della turbolenza; tanto in un caso quanto nell'altro, la scienza e la tecnologia dell'epoca non lo permettevano. Le teorie moderne sulla turbolenza trovano corrispondenza nell'anticipazione un po' vaga di Poincaré, ma vanno ben più lontano, dato che permettono di stabilire la prevedibilità dei fenomeni. I meteorologi sanno ora descrivere configurazioni di migliore o peggiore prevedibilità e sanno anche quantificare questa prevedibilità, cosa che era del tutto impensabile qualche anno fa.

Inoltre, le idee moderne sulla prevedibilità possono estendersi a molteplici dominî che non appartengono solo alla fisica, ma anche alla chimica, all'ecologia, all'economia ecc. Anche se in questi dominî non si può dimostrare la dipendenza sensibile dalle condizioni iniziali, la sua probabile esistenza è di grande importanza filosofica. Questo dovrebbe far riflettere i sostenitori del determinismo storico. Anche ammettendo che la storia abbia leggi deterministiche, il suo corso può comunque risultare imprevedibile: le vie del Signore sono infinite...

Questo attrattore strano a tre dimensioni è stato osservato, presso l'Università del Texas ad Austin, da Jean-Claude Roux, Reuben Simoyi e Harry Swinney. Questi ricercatori hanno analizzato la reazione di Belousov-Žabotinskij, nella quale una soluzione di acido bromico ossida l'acido malonico in presenza di ioni metallici, che fungono da catalizzatori, e di altri composti necessari per controllare la reazione, e hanno riportato in uno spazio tridimensionale le concentrazioni dei reagenti. In alto è rappresentata una proiezione bidimensionale di questo attrattore a tre dimensioni. In basso sono mostrate diverse sezioni dell'attrattore. Si può così osservare che condizioni notevolmente simili producono effetti alquanto diversi; l'allungamento e il restringimento delle sezioni successive sono provocati dal fenomeno della dipendenza sensibile dalle condizioni iniziali.

Il dito di Dio o la mano del diavolo?

Che cosa abbiamo detto? Vogliamo sostenere che l'imprevedibilità della storia umana, inerente alle leggi fisiche, dà a Dio l'occasione d'intervenire nei nostri affari? O al diavolo? È un'idea che è stata avanzata ai tempi della scoperta dell'indeterminismo quantistico. Alcuni diranno che si tratta di una questione metafisica che il fisico non ha il diritto di porre. Rendiamo dunque ai teologi ciò che a loro appartiene e immaginiamo un demone fittizio che voglia interferire nei nostri affari attraverso impercettibili aggiustamenti dell'universo. Potrebbe farlo? Facilmente, come ora vedremo.

Quando abbiamo preso in considerazione il gas di sfere dure, abbiamo visto che un piccolo cambiamento delle condizioni iniziali potrebbe provocare, dopo un certo numero di collisioni, un cambiamento qualitativo: due sfere che avrebbero dovuto entrare in collisione non lo fanno più. Émile Borel, matematico (e politico) francese, aveva già stimato il tempo necessario a questo cambiamento qualitativo. Il fisico Michael Berry dell'Università di Bristol ha ripreso il calcolo identificando un gas comune con un gas di sfere dure (operazione un po' disinvolta, ma accettabile per il problema che stiamo esaminando). Berry immagina che le forze che agiscono sul gas vengano modificate eliminando l'attrazione gravitazionale di un elettrone posto al limite dell'universo conosciuto, e si chiede quanto tempo bisognerà attendere perché abbia luogo una modificazione qualitativa negli urti molecolari all'interno del gas. La risposta è la seguente: questo tempo è uguale al tempo necessario perché una molecola abbia 50 collisioni con altre molecole, cioè un'infima frazione di secondo.

Senza fatica, il nostro demone è riuscito a modificare la struttura microscopica di un gas, ma si tratta ancora di dettagli impercettibili per i nostri sensi grossolani. In che modo arriveremo a cambiamenti importanti? La turbolenza viene a questo punto in aiuto del demone, dato che può amplificare una modificazione piccolissima (una fluttuazione microscopica) e renderla macroscopica in un tempo abbastanza breve. Secondo una stima fatta utilizzando nozioni comunemente accettate sulla turbolenza, ci vuole circa un minuto perché fluttuazioni microscopiche si trasformino in cambiamenti macroscopici nel tipo di turbolenza che si produce sopra un termosifone caldo.

Si può anche stimare il tempo necessario alle fluttuazioni per propagarsi dalle

piccole dimensioni alle dimensioni superiori. Si può così calcolare che ci vuole un giorno perché cambiamenti su distanze dell'ordine di un centimetro si trasformino in cambiamenti su distanze dell'ordine dei dieci chilometri. Per tornare al demone che interferirebbe negli affari umani, può dunque riuscire in un giorno a perturbare un temporale locale. Quanto tempo gli ci vorrebbe per cambiare completamente la struttura delle condizioni meteorologiche sulla Terra? I meteorologi, soprattutto gli statunitensi Charles Leith e Robert Kraichman, hanno studiato la questione e le stime attuali sono dell'ordine di una o due settimane.

Stando così le cose, il minimo cambiamento nell'universo può ripercuotersi in cambiamenti così grandi da cambiare letteralmente la faccia del mondo in 15 giorni. Il mondo in cui viviamo è forse deterministico e noi possiamo certamente prevederne in una certa misura l'evoluzione, ma, sempre in una certa misura, questa evoluzione è imprevedibile; e questa imprevedibilità possiamo oggi stimarla, calcolarla.

numero 222
febbraio 1987
anno XX
volume XXXVIII

LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

Il caos

Vi è ordine nel caos: l'aleatorietà possiede una forma geometrica soggiacente; il caos pone limitazioni fondamentali alla prevedibilità, ma suggerisce relazioni causali dove prima non parevano possibili

di James P. Crutchfield, J. Doyne Farmer, Norman H. Packard e Robert S. Shaw

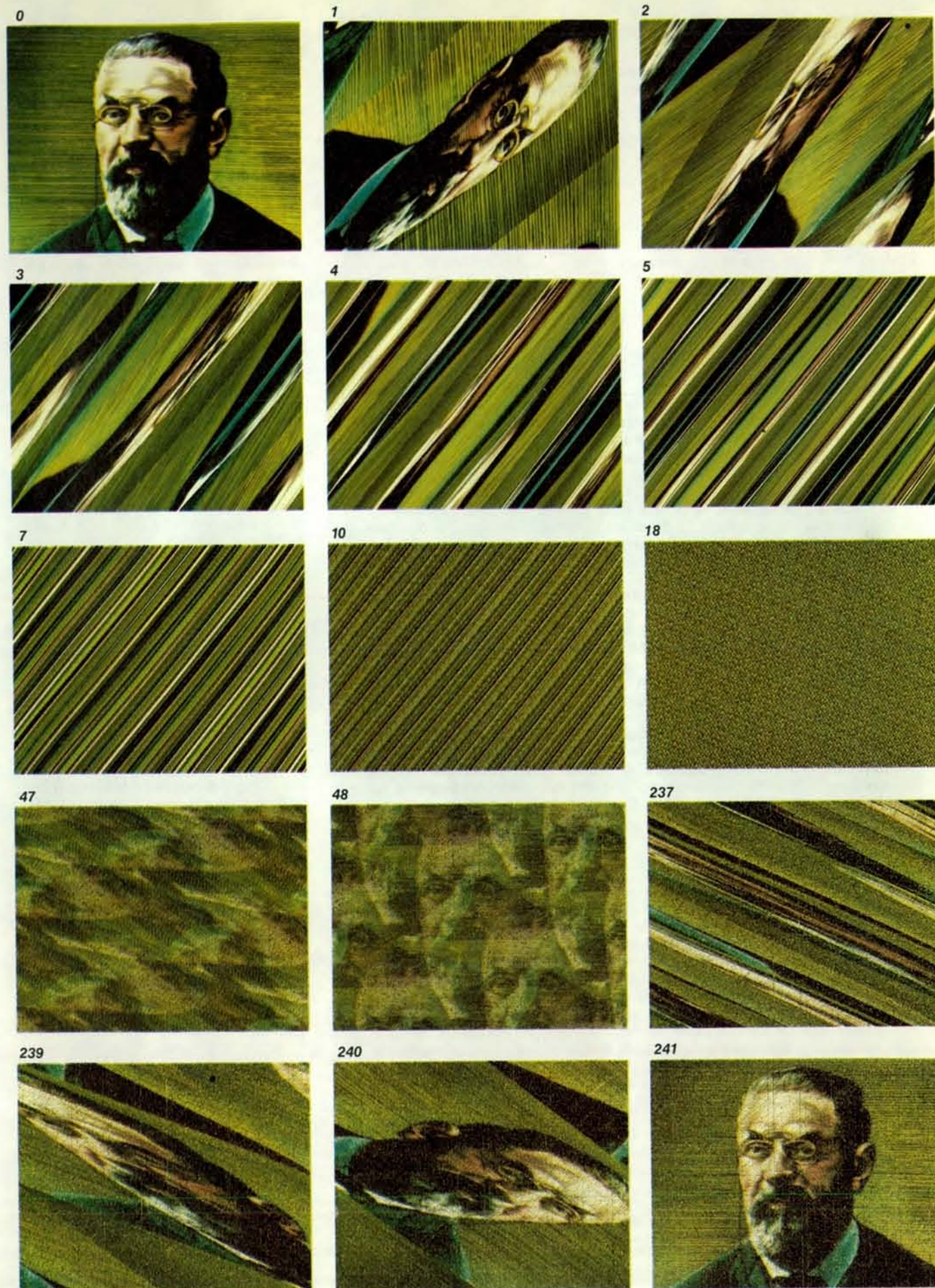
La grande forza della scienza sta nella sua capacità di collegare cause ed effetti. Sulla base delle leggi della gravitazione, per esempio, è possibile prevedere le eclissi con migliaia di anni di anticipo. Altri fenomeni naturali, però, non sono altrettanto prevedibili. Il moto dell'atmosfera segue le leggi della fisica al pari del moto dei pianeti; tuttavia le previsioni del tempo vengono ancora espresse in termini probabilistici. L'evoluzione delle condizioni meteorologiche, lo scorrere di un ruscello, il rotolare di un dado presentano aspetti imprevedibili. Poiché in questi fenomeni non esiste una chiara relazione tra causa ed effetto, si dice che contengono elementi aleatori. Eppure fino a poco tempo fa non vi erano validi motivi per dubitare che in linea di principio si potesse giungere a una prevedibilità rigorosa. Si riteneva che bastasse raccogliere ed elaborare una quantità sufficiente di informazione.

Questa opinione è stata modificata da una scoperta sorprendente. Semplici sistemi deterministici, anche costituiti da pochi elementi, possono manifestare un comportamento aleatorio. Questa aleatorietà è di natura essenziale e non scompare se si raccolgono ulteriori informazioni. A questo genere di aleatorietà si è dato il nome di caos.

Un paradosso apparente è che il caos è deterministico, cioè è generato da regole fisse che, di per sé, non contengono alcun elemento casuale. In linea di principio il futuro è determinato completamente dal passato, ma in pratica le piccole indeterminazioni vengono amplificate; quindi, benché il comportamento sia prevedibile a breve scadenza, alla lunga risulta imprevedibile. Nel caos vi è ordine: soggiacenti al comportamento caotico vi sono eleganti forme geometriche che creano l'aleatorietà così come il cartaio mescola un mazzo di carte o un cuoco mescola l'impasto di un dolce.

La scoperta del caos ha creato un paradigma nuovo fra i modelli scientifici. Da una parte comporta l'esistenza di nuove limitazioni fondamentali alla nostra capacità di compiere previsioni; dall'altra, il determinismo inerente al caos implica che molti fenomeni aleatori sono più prevedibili di quanto si pensasse. Informazioni apparentemente aleatorie raccolte in passato (e archiviate perché ritenute troppo complicate) oggi possono essere spiegate in termini di leggi semplici. Il caos ci consente di scoprire l'ordine in sistemi diversissimi fra loro come l'atmosfera, un rubinetto che gocciola e il cuore. Ne è conseguita una rivoluzione che sta coinvolgendo molte branche diverse della scienza.

Che origine ha il comportamento caotico? Il moto browniano è un esempio classico di aleatorietà. Osservando al microscopio un granello di polvere in una bacinella d'acqua, lo si vede muoversi con un tremolio continuo e capriccioso. Ciò è dovuto alle molecole d'acqua circostanti che, mosse dall'agitazione termica, bombardano la particella di polvere. Poiché le molecole d'acqua sono invisibili e sono numerosissime, il movimento preciso della particella di polvere è affatto imprevedibile. Qui la rete delle azioni causali tra le varie sottounità può essere così intricata da rendere affatto aleatoria la configurazione del moto risultante.

Il caos deriva dall'operazione geometrica dello stiramento. Su un ritratto del matematico francese Henri Poincaré, il fondatore della teoria dei sistemi dinamici, ne è illustrato l'effetto. L'immagine di partenza (*in alto a sinistra*) è stata discretizzata per consentire a un calcolatore di effettuare lo stiramento. Una semplice trasformazione matematica stira l'immagine diagonalmente, come se fosse dipinta su un foglio di gomma. Quando il foglio esce dal riquadro, viene tagliato e reinserito dall'altra parte, come si vede nell'immagine *1*. (Il numero sopra ciascuna immagine indica quante volte è stata effettuata la trasformazione.) La ripetizione dell'operazione ha per effetto di scompaginare la faccia (*immagini* 2-4). Il risultato finale è una combinazione casuale di colori, che si presenta come uno strato omogeneo di verde (*immagini* 10 *e* 18). Talvolta accade che alcuni punti tornino in prossimità delle loro posizioni iniziali, provocando una breve ricomparsa dell'immagine originale (*immagini* 47-48, 239-241). La trasformazione illustrata è particolare in quanto il fenomeno della «ricorrenza di Poincaré» (come viene chiamata in meccanica statistica) si presenta molto più spesso del solito; in generale in una trasformazione caotica generica la ricorrenza è rarissima e si presenta forse una volta sola nel corso della vita dell'universo. In presenza di fluttazioni di fondo di entità qualunque, l'intervallo tra una ricorrenza e l'altra è di solito così lungo che tutta l'informazione relativa all'immagine originale va perduta.

Il tipo di caos di cui parleremo non ha bisogno di un gran numero di sottounità o di azioni invisibili. La presenza di un comportamento aleatorio in sistemi semplicissimi giustifica un riesame delle fonti dell'aleatorietà anche in grandi sistemi, come quello meteorologico.

Che cosa rende il moto dell'atmosfera tanto più difficile da prevedere del moto del sistema solare? Entrambi i sistemi sono costituiti da molte componenti e sono retti dalla seconda legge della dinamica, $F = ma$, che può essere considerata una semplice ricetta per prevedere il futuro. Se le forze F agenti su una massa m assegnata sono note, lo è anche l'accelerazione a. Dalle regole dell'analisi segue allora che se la posizione e la velocità di un oggetto possono essere misurate a un istante dato, esse restano determinate per sempre. Si tratta di un principio molto potente, tanto che nel Settecento il grande matematico francese Simon de Laplace affermò che, date la posizione e la velocità di tutte le particelle dell'universo, egli sarebbe stato capace di prevederne l'evoluzione futura per l'eternità. Benché molte evidenti difficoltà pratiche si oppongano alla realizzazione del programma di Laplace, per oltre un secolo sembrò che nulla ne infirmasse la legittimità, almeno in linea di principio. L'applicazione letterale dell'affermazione di Laplace al comportamento umano portò alla conclusione filosofica che questo comportamento fosse completamente predeterminato e che il libero arbitrio non esistesse.

La scienza del Novecento ha visto il crollo del determinismo di Laplace e per due ragioni molto diverse. La prima è la meccanica quantistica: un dogma fondamentale di questa teoria è il principio d'indeterminazione di Heisenberg, secondo il quale vi è una limitazione essenziale alla precisione con cui possono essere misurate la posizione e la velocità di una particella. Questa indeterminazione dà una buona spiegazione di certi fenomeni aleatori, come la disintegrazione radioattiva. Il nucleo è così piccolo che il principio d'indeterminazione pone una limitazione fondamentale alla conoscenza del suo moto ed è impossibile raccogliere informazioni sufficienti a prevedere il momento della sua disintegrazione.

Su scala più grande, però, la fonte dell'imprevedibilità dev'essere cercata altrove. Alcuni fenomeni macroscopici sono prevedibili e altri no, e questa distinzione non ha nulla a che fare con la meccanica quantistica. La traiettoria di una palla, per esempio, è intrinsecamente prevedibile, e un giocatore sfrutta istintivamente questo fatto. Invece la traiettoria di un palloncino spinto dall'aria che ne esce impetuosamente non è prevedibile; il palloncino caracolla e muta rotta capricciosamente in istanti e in punti che è assolutamente impossibile prevedere. Il palloncino segue le leggi di Newton esattamente come la palla: allora, perché è tanto più difficile prevederne il comportamento?

L'esempio classico di questa dicotomia è il moto dei fluidi. In certe condizioni il moto di un fluido è laminare (cioè uniforme, stazionario e regolare) e facile da prevedere a partire dalle equazioni. In altre condizioni il moto del fluido è turbolento (cioè non uniforme, non stazionario e irregolare) e quindi difficile da prevedere. Il passaggio dal comportamento laminare a quello turbolento è ben noto a chiunque, viaggiando in aereo, sia passato improvvisamente da una zona di calma a una di tempesta. Qual è la causa della differenza essenziale tra il moto laminare e quello turbolento?

Per capire appieno perché si tratti di un vero e proprio enigma, pensiamo a un ruscello di montagna. L'acqua turbina e schizza come se avesse una mente propria, spostandosi prima di qua e poi di là. Tuttavia i massi sono ben piantati nel letto del ruscello e i suoi tributari gli forniscono acqua a un tasso quasi costante. Da dove nasce allora l'aleatorietà del moto dell'acqua?

Al fisico sovietico Lev D. Landau, scomparso di recente, è attribuita una spiegazione del moto aleatorio dei fluidi che ha tenuto il campo per molti anni: il moto di un fluido turbolento conterrebbe molte oscillazioni diverse e indipendenti. Via via che il fluido si muove più in fretta, e diventa perciò più turbolento, le oscillazioni entrano nel moto una alla volta. Benché le singole oscillazioni possano essere semplici, la complessità del moto combinato rende impossibile prevedere il flusso complessivo.

La teoria di Landau, tuttavia, è stata confutata. Il comportamento aleatorio si presenta anche in sistemi semplicissimi, senza bisogno di complicazioni o indeterminazioni. Il matematico francese Henri Poincaré se ne rese conto all'inizio del secolo, quando osservò che fenomeni imprevedibili e «fortuiti» possono presentarsi nei sistemi in cui una piccola variazione nel presente provoca una variazione molto più ampia nel futuro. Il concetto si chiarisce se si pensa a un sasso posto in equilibrio sulla cima di un monte. Una spintarella in un verso o in un altro basta a farlo rotolare giù lungo percorsi molto diversi. Il sasso è sensibile a piccole azioni solo quando è sulla cima del monte; i sistemi caotici invece sono sensibili in ogni punto del loro moto.

Un semplice esempio servirà a illustrare quanto certi sistemi fisici siano sensibili alle azioni esterne. Consideriamo una partita a biliardo; idealizzandola un po' si può supporre che le palle si muovano sul tavolo e si urtino con una perdita di energia trascurabile. Con un solo colpo il giocatore provoca fra tutte le palle una prolungata successione di collisioni. Naturalmente il giocatore desidera conoscere gli effetti del colpo. Se possedesse un controllo perfetto del proprio colpo, per quanto tempo potrebbe dimostrarsi valida la previsione della traiettoria della palla colpita dalla stecca? Ebbene, se il giocatore ignorasse effetti anche minimi, per esempio l'attrazione gravitazionale di un elettrone situato ai margini della galassia, la sua previsione sarebbe errata già dopo un solo minuto!

Questo cospicuo aumento dell'indeterminazione deriva dal fatto che le palle sono curve e che a ciascuna collisione vengono amplificate le piccole differenze nel punto di contatto. Quest'amplificazione è esponenziale: viene composta a ogni collisione, come accade per le generazioni successive dei batteri quando spazio e cibo siano illimitati. Qualunque effetto, per quanto piccolo, acquista rapidamente proporzioni macroscopiche. Questa è una delle caratteristiche fondamentali del caos.

L'amplificazione esponenziale degli errori dovuta alla dinamica del caos costituisce la seconda ragione dell'insuccesso di Laplace. La meccanica quantistica comporta che le misure iniziali siano sempre imprecise e il caos fa sì che le imprecisioni superino rapidamente la capacità di fare previsioni. Senza il caos Laplace avrebbe potuto sperare che gli errori restassero limitati, o almeno crescessero abbastanza lentamente da consentirgli di fare previsioni a lungo termine. In presenza del caos le previsioni sono invece destinate a raggiungere ben presto un'imprecisione madornale.

Il più ampio quadro concettuale dal quale il caos emerge è la cosiddetta teoria dei sistemi dinamici. Un sistema dinamico si compone di due parti: le caratteristiche del suo stato (cioè le informazioni essenziali sul sistema) e la dinamica (una regola che descrive l'evoluzione dello stato nel tempo). L'evoluzione può essere visualizzata in uno spazio degli stati, una costruzione astratta le cui coordinate sono le componenti dello stato. In generale le coordinate dello spazio degli stati variano secondo il contesto; per un sistema meccanico potrebbero essere la posizione e la velocità, ma per un modello ecologico potrebbero essere le popolazioni delle diverse specie.

Un buon esempio di sistema dinamico è offerto dal pendolo semplice. Per descrivere completamente il suo moto bastano due variabili: la posizione e la velocità. Lo stato è quindi un punto del piano le cui coordinate sono posizione e velocità. Le leggi di Newton forniscono una regola, che assume la forma matematica di un'equazione differenziale, per descrivere l'evoluzione dello stato. Quando il pendolo oscilla avanti e indietro, il suo stato si sposta lungo una curva del piano, detta «orbita». Nel caso ideale di un pendolo privo di attrito, l'orbita è una curva chiusa; altrimenti è una curva a spirale che converge verso un punto corrispondente allo stato di quiete del pendolo.

L'evoluzione temporale di un sistema dinamico può avvenire sia in tempo continuo sia in tempo discreto. Nel primo caso si parla di flusso, nel secondo di applicazione. Un pendolo si muove con continuità da uno stato a un altro e quindi viene descritto da un flusso in tempo continuo. Il numero degli insetti che nascono ogni anno in una data zona e l'intervallo temporale tra le gocce di un rubinetto che gocciola vengono invece descritti in modo più naturale da un'applicazione in tempo discreto.

Per vedere come un sistema si evolva a partire da un dato stato iniziale si può usare la dinamica (cioè le equazioni del moto) per effettuare uno spostamento incrementale lungo un'orbita. Questo metodo per ricavare il comportamento del sistema richiede un impegno di calcolo proporzionale all'intervallo di tempo in cui si desidera seguire l'orbita. Per certi sistemi semplici, come il pendolo senza attrito, le equazioni del moto possono talora possedere una soluzione in forma chiusa, che è una formula che esprime qualunque stato futuro in termini dello stato iniziale. Una soluzione in forma chiusa rappresenta una scorciatoia, un algoritmo più semplice, che consente di prevedere il futuro solo in base allo stato iniziale e all'istante finale, senza bisogno di passare per gli stati intermedi. Con una soluzione di questo tipo l'impegno algoritmico necessario per seguire il moto del sistema è grosso modo indipendente dal tempo desiderato. Date le equazioni del moto dei pianeti e della Luna, nonché le posizioni e le velocità della Terra e della Luna, si possono prevedere, per esempio, le eclissi con anni di anticipo.

I successi ottenuti agli inizi della fisica nel trovare soluzioni in forma chiusa per svariati sistemi semplici fecero nascere la speranza che soluzioni di questo genere esistessero per qualunque sistema meccanico. Oggi si sa che ciò, purtroppo, in

Laplace, 1776

«Lo stato attuale del sistema della natura consegue evidentemente da quello che era all'istante precedente e se noi immaginassimo un'intelligenza che a un istante dato comprendesse tutte le relazioni fra le entità di questo universo, essa potrebbe conoscere le rispettive posizioni, i moti e le disposizioni generali di tutte quelle entità in qualunque istante del passato o del futuro.

«L'astronomia fisica, la branca della scienza che rende il massimo onore alla mente umana, ci dà un'idea, ancorché imperfetta, di quale sarebbe questa intelligenza. La semplicità della legge che regola il moto dei corpi celesti e le relazioni fra le loro masse e distanze ci consentono di seguire i loro moti con l'analisi fino a un certo punto; e per poter determinare lo stato del sistema di questi grandi corpi nei secoli passati o in quelli futuri è sufficiente al matematico conoscere tramite l'osservazione la loro posizione e la loro velocità in un istante qualunque. L'uomo deve questo vantaggio alla potenza degli strumenti impiegati e al piccolo numero di relazioni che entrano nei suo calcoli. Ma l'ignoranza delle diverse cause che concorrono alla formazione degli eventi come pure la loro complessità, insieme con l'imperfezione dell'analisi, ci impediscono di conseguire la stessa certezza rispetto alla grande maggioranza dei fenomeni. Vi sono quindi cose che per noi sono incerte, cose più o meno probabili, e noi cerchiamo di rimediare all'impossibilità di conoscerle determinando i loro diversi gradi di verosimiglianza. Accade così che alla debolezza della mente umana si debba una delle più fini e ingegnose fra le teorie matematiche, la scienza del caso o della probabilità.»

Poincaré, 1903

«Una causa piccolissima che sfugga alla nostra attenzione determina un effetto considerevole che non possiamo mancar di vedere, e allora diciamo che l'effetto è dovuto al caso. Se conoscessimo esattamente le leggi della natura e la situazione dell'universo all'istante iniziale, potremmo prevedere esattamente la situazione dello stesso universo in un istante successivo. Ma se pure accadesse che le leggi naturali non avessero più alcun segreto per noi, anche in tal caso potremmo conoscere la situazione iniziale solo *approssimativamente*. Se questo ci permettesse di prevedere la situazione successiva con la stessa *approssimazione*, non ci occorrerebbe di più e dovremmo dire che il fenomeno è stato previsto, che è governato da leggi. Ma non sempre è così; può accadere che piccole differenze nelle condizioni iniziali ne producano di grandissime nei fenomeni finali. Un piccolo errore nelle prime produce un errore enorme nei secondi. La previsione diviene impossibile e si ha un fenomeno fortuito.»

Le opinioni di due luminari sul caso e sulla probabilità messe a confronto. Il matematico francese Pierre Simon de Laplace sosteneva che le leggi della natura implicano un rigido determinismo e una prevedibilità totale, benché le imperfezioni delle osservazioni rendano necessario introdurre la teoria delle probabilità. Nella citazione di Poincaré è prefigurata l'opinione odierna secondo al quale indeterminazioni arbitrariamente piccole sullo stato del sistema possono amplificarsi con il passare del tempo, rendendo impossibili le previsioni sul futuro lontano.

Lo spazio degli stati è un concetto utile per visualizzare il comportamento di un sistema dinamico. Si tratta di uno spazio astratto le cui coordinate sono i gradi di libertà del moto del sistema. Il moto di un pendolo (*in alto*), per esempio, è determinato dalla posizione e dalla velocità iniziali. Il suo stato è quindi un punto di un piano le cui coordinate sono posizione e velocità (*in basso*). Quando il pendolo compie le sue oscillazioni il punto descrive un'«orbita», cioè un cammino, nello spazio degli stati. Per un pendolo ideale, senza attrito, l'orbita è una curva chiusa (*in basso a sinistra*); se c'è attrito, l'orbita è una spirale che tende a un punto (*in basso a destra*).

generale non è vero. Il comportamento imprevedibile dei sistemi dinamici caotici non può essere espresso mediante una soluzione in forma chiusa e di conseguenza non esiste alcuna scorciatoia per prevedere tale comportamento.

Nondimeno lo spazio degli stati costituisce uno strumento potente per descrivere il comportamento dei sistemi caotici. Lo spazio degli stati è utile perché consente di rappresentare il comportamento in forma geometrica. Per esempio, un pendolo soggetto ad attrito prima o poi si arresta e ciò nello spazio degli stati corrisponde a un'orbita che tende verso un punto. Il punto non si muove - è un punto fisso - e, poiché attrae le orbite vicine, lo si chiama attrattore. Se si dà una debole spinta a un pendolo, esso torna allo stesso attrattore fisso. Tutti i sistemi che col tempo pervengono a uno stato di quiete possono essere caratterizzati da un punto fisso nello spazio degli stati. Questo è un caso particolare di un fenomeno generalissimo, che si presenta quando le perdite, dovute per esempio all'attrito o alla viscosità, fanno sì che le orbite siano attratte verso una regione più piccola dello spazio degli stati avente un numero inferiore di dimensioni. Anche queste regioni sono chiamate attrattori. In termini grossolani, un attrattore è ciò in cui si stabilizza o verso cui è attratto il comportamento di un sistema.

Alcuni sistemi alla lunga non tendono alla quiete e passano invece ciclicamente per una successione periodica di stati. Un esempio è costituito dall'orologio a pendolo, in cui l'energia perduta per gli attriti viene reintegrata da una molla o da un peso. Il pendolo continua a ripetere lo stesso moto. Nello spazio degli stati un moto siffatto corrisponde a un'orbita periodica o ciclo. Indipendentemente da come un pendolo venga posto in oscillazione, il ciclo che rappresenta il limite cui esso tende a lungo termine è sempre lo stesso. Gli attrattori di questo tipo si chiamano pertanto cicli limite. Un altro sistema ben noto che ha come attrattore un ciclo limite è il cuore.

Un sistema può possedere più attrattori. In tal caso condizioni iniziali diverse possono portare ad attrattori diversi. L'insieme dei punti che si evolvono verso un dato attrattore costituisce il suo bacino di attrazione. L'orologio a pendolo possiede due di questi bacini: se lo spostamento del pendolo dalla posizione di quiete è piccolo, esso torna in questa posizione; se invece lo spostamento è grande, l'orologio si avvia e il pendolo si mette a oscillare in modo stabile.

La forma di attrattore immediatamente più complicata è il toro, che somiglia alla superficie di una ciambella. Questa forma descrive un moto costituito da due oscillazioni indipendenti, moto talvolta detto quasiperiodico. (Se ne possono costruire esempi fisici mediante oscillatori elettrici pilotati.) Nello spazio degli stati l'orbita si avvolge intorno al toro; una delle due frequenze è determinata dalla velocità con cui l'orbita gira intorno al toro nella dimensione minore, mentre l'altra dipende dalla velocità con cui l'orbita si sposta lungo la dimensione maggiore. Gli attrattori possono essere anche tori con un numero più elevato di dimensioni e che rappresentano una combinazione di più di due oscillazioni.

La caratteristica più importante del moto quasiperiodico è che, nonostante la sua complessità, esso è prevedibile. Anche se l'orbita non si ripete mai esattamente, come accade se le frequenze del moto non hanno un divisore comune, il moto resta regolare. Le orbite che cominciano vicine tra loro sul toro rimangono vicine tra loro e la prevedibilità a lungo termine è assicurata.

Gli attrattori sono forme geometriche che caratterizzano il comportamento a lungo termine nello spazio degli stati. In altri termini, un attrattore è ciò verso cui si stabilizza, o è attratto, il comportamento di un sistema. Nel caso raffigurato gli attrattori sono in blu mentre gli stati iniziali sono in rosso. Le traiettorie (*in verde*), partendo dagli stati iniziali, finiscono con l'avvicinarsi agli attrattori. Il tipo di attrattore più semplice è un punto fisso (*in alto a sinistra*). Questo attrattore corrisponde a un pendolo soggetto ad attrito: il pendolo ritorna sempre nella stessa posizione di quiete, indipendentemente da come abbia cominciato a oscillare (*si veda la parte a destra dell'illustrazione della pagina precedente*). L'attrattore immediatamente più complicato è un ciclo limite (*in alto al centro*), che nello spazio degli stati è una curva chiusa. Un ciclo limite descrive oscillazioni stabili, come il moto di un orologio a pendolo o il battito cardiaco. Le oscillazioni composte, caratteristiche del comportamento quasiperiodico, corrispondono a un attrattore a forma di toro (*in alto a destra*). Tutti e tre questi attrattori sono prevedibili: il loro comportamento può essere previsto con una precisione arbitraria. Gli attrattori caotici, invece, corrispondono a moti imprevedibili e hanno una forma geometrica più complicata. Nella parte inferiore dell'immagine sono illustrati tre esempi di attrattori caotici; da sinistra a destra, essi sono stati scoperti rispettivamente da Edward N. Lorenz, da Otto E. Rössler e da uno degli autori (Shaw). Le immagini sono state ottenute ricorrendo a semplici sistemi di equazioni differenziali aventi uno spazio degli stati tridimensionale.

Fino a tempi piuttosto recenti gli unici attrattori conosciuti erano i punti, i cicli limite e i tori. Nel 1963 Edward N. Lorenz, del Massachusetts Institute of Technology, scoprì un esempio concreto di sistema a poche dimensioni che ha un comportamento complesso. Spinto dal desiderio di comprendere l'imprevedibilità delle condizioni meteorologiche, Lorenz partì dalle equazioni del moto di un fluido (l'atmosfera può essere considerata un fluido) e semplificandole ottenne un sistema dotato di tre soli gradi di libertà. Il sistema, però, si comportava in un modo evidentemente aleatorio che non poteva essere caratterizzato adeguatamente da nessuno dei tre attrattori noti a quel tempo. L'attrattore che osservò, e che ora viene chiamato attrattore di Lorenz, fu il primo esempio di attrattore caotico, o attrattore strano.

Servendosi di un calcolatore digitale per simulare il proprio semplice modello, Lorenz chiarì il meccanismo fondamentale che dava luogo all'aleatorietà osservata: le perturbazioni microscopiche vengono amplificate fino a interferire con il comportamento macroscopico. Due orbite corrispondenti a condizioni iniziali prossime divergono con velocità esponenziale e quindi restano vicine tra loro soltanto per breve tempo. Per gli attrattori non caotici la situazione è qualitativamente diversa, poiché le orbite vicine restano vicine l'una all'altra, gli errori piccoli restano limitati e il comportamento è prevedibile.

La chiave per interpretare il comportamento caotico sta nella comprensione di una semplice operazione di stiramento e piegatura, che ha luogo nello spazio degli stati. La divergenza esponenziale è un fenomeno locale: dal momento che la dimensione (*size*) degli attrattori è finita, due orbite situate su un attrattore caotico non possono continuare a divergere esponenzialmente per sempre. Ne segue che l'attrattore deve ripiegarsi su se stesso. Benché le orbite divergano e seguano strade sempre più diverse, prima o poi devono passare di nuovo una accanto all'altra. Le orbite situate su un attrattore caotico vengono mescolate da questo processo più o meno come un mazzo di carte viene mescolato dal car-

Un attrattore caotico ha una struttura molto più complicata di uno prevedibile come un punto, un ciclo limite o un toro. Se molto ingrandito, un attrattore caotico si presenta come una superficie non regolare e contenente pieghe. L'immagine presenta i passaggi che portano alla costruzione di un attrattore caotico nel caso più semplice: quello di Rössler (*in basso*). Le traiettorie vicine sull'oggetto devono «stirarsi», cioè divergere, esponenzialmente (*in alto*); nel caso illustrato la distanza fra traiettorie vicine diventa circa il doppio. Inoltre, per restare compatto, l'oggetto deve «ripiegarsi» su se stesso (*al centro*): la superficie si piega su se stessa finché le due estremità s'incontrano. Osservato in molti sistemi, dal moto dei fluidi alle reazioni chimiche, l'attrattore di Rössler illustra la massima di Einstein che la natura predilige le forme semplici.

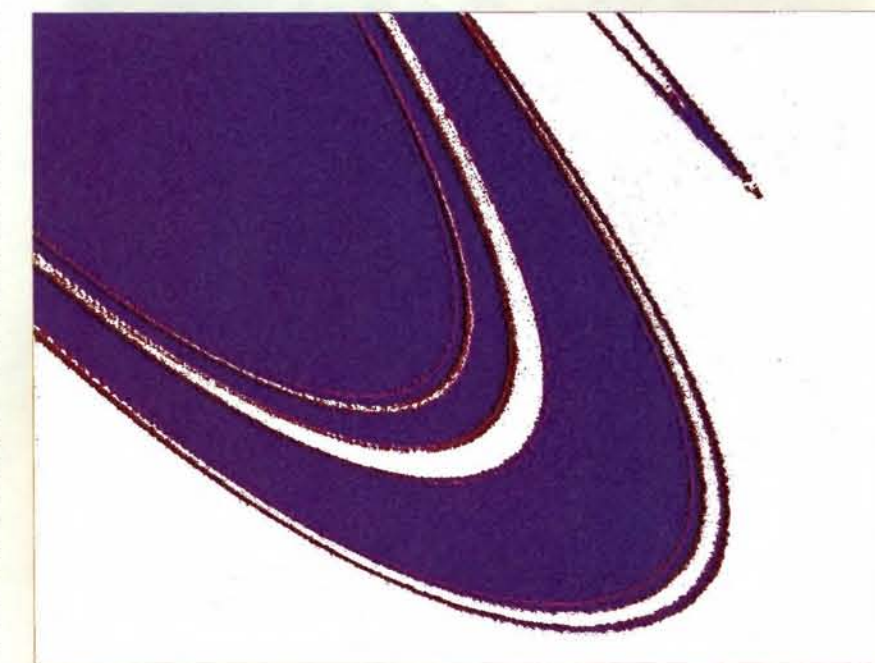
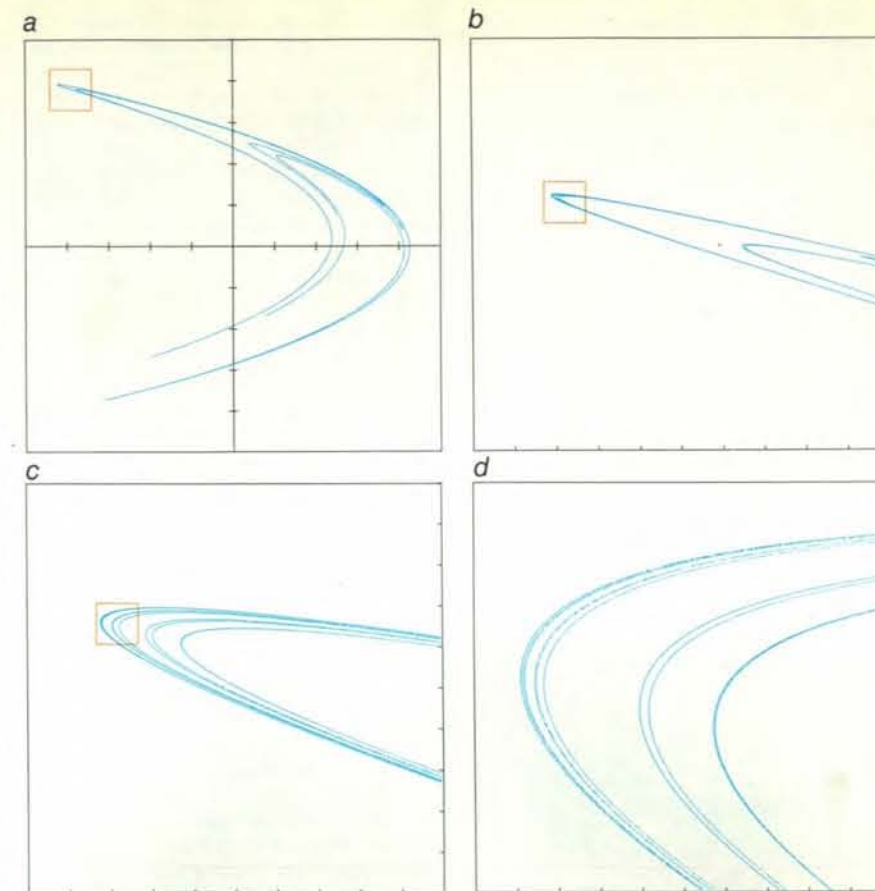
taio. L'aleatorietà delle orbite caotiche è una conseguenza di questo processo di mescolamento. Il processo di stiramento e piegatura avviene più volte e produce pieghe dentro altre pieghe, all'infinito. In altre parole, un attrattore caotico è un frattale, cioè un oggetto che rivela particolari sempre più numerosi via via che viene ingrandito (*si veda l'illustrazione nella pagina a fronte*).

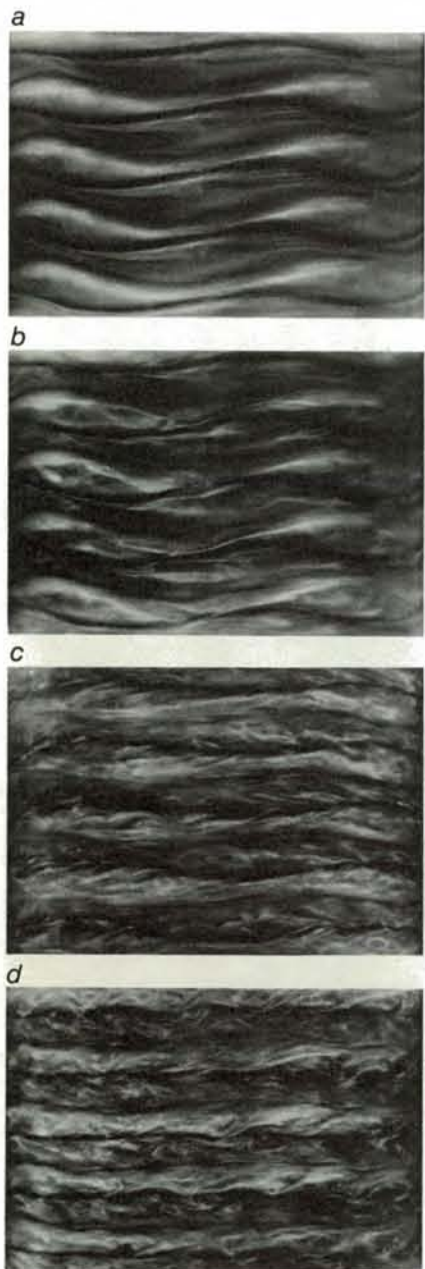
Il caos mescola le orbite nello spazio degli stati esattamente come un fornaio impasta il pane. Ci si può immaginare ciò che accade alle traiettorie vicine su un attrattore caotico versando nella pasta una goccia di colorante blu. L'impastatura è una combinazione di due azioni: lo stendimento della pasta, che fa diffondere il colorante, e il ripiegamento della pasta su se stessa. Dapprima la chiazza di colorante semplicemente si allunga, ma poi viene ripiegata e dopo un tempo piuttosto lungo si trova stirata e ripiegata molte volte. Osservandolo da vicino, si vede che l'impasto consiste in molti strati, alternativamente blu e bianchi. Già dopo 20 passaggi la lunghezza iniziale della macchia è aumentata più di un milione di volte e il suo spessore si è assottigliato fino a dimensioni molecolari. Il colorante blu è completamente mescolato con l'impasto. Il caos agisce allo stesso modo, ma invece di mescolare pasta mescola lo spazio degli stati. Ispirandosi a questa immagine del mescolamento, Otto E. Rössler, dell'Università di Tübingen, ha ideato l'esempio più semplice di attrattore caotico in un flusso (*si veda l'illustrazione a pagina 15*).

La divergenza delle traiettorie vicine è la ragione di fondo per la quale il caos porta all'imprevedibilità. Una misurazione perfetta corrisponderebbe a un punto dello spazio degli stati, ma tutte le misurazioni reali sono imprecise e generano un alone di indeterminazione. Lo stato effettivo può essere in qualunque punto interno a questo alone. Come si vede qui nel caso dell'attrattore di Lorenz, l'imprecisione della misura iniziale è rappresentata da 10 000 puntini rossi, che all'inizio sono così vicini tra loro da essere indistinguibili. Via via che ciascun punto si muove per effetto delle equazioni, l'alone viene stirato e forma un lungo filo sottile, che poi si piega su se stesso molte volte, fino a quando i punti risultano diffusi su tutto l'attrattore. Ora la previsione è impossibile: lo stato finale può essere in qualunque punto dell'attrattore. Per un attrattore prevedibile, invece, tutti gli stati finali rimangono vicini tra loro. I numeri sopra l'illustrazione sono in unità di 1/200 di secondo.

Quando si compiono osservazioni su un sistema fisico, è impossibile determinare esattamente lo stato del sistema a causa degli inevitabili errori di misurazione. Quindi lo stato del sistema non è situato in un unico punto, bensì all'interno di una piccola regione dello spazio degli stati. Benché sia l'indeterminazione quantistica che impone le dimensioni ultime di questa regione, in pratica vari tipi di rumore limitano la precisione delle misurazioni introducendo errori sostanzialmente più grandi. La piccola regione determinata da una misurazione corrisponde alla chiazza di colorante blu nell'impasto.

Quando si effettua una misurazione collocando il sistema in una piccola regione dello spazio degli stati, si ricava una certa quantità di informazione sul sistema. Quanto più accurata è la misurazione, tanto maggiore è la conoscenza che l'osservatore ricava sullo stato del sistema. Viceversa, quanto più ampia è la regione tanto più incerto è l'osservatore. Poiché nei sistemi non caotici i punti vicini rimangono vicini durante l'evoluzione temporale, una misurazione fornisce una certa quantità d'informazione che si conserva nel tempo. Questo è proprio il senso in cui questi sistemi sono prevedibili: le misurazioni iniziali contengono informazioni che possono essere sfruttate per prevederne il comportamento futuro. In altri termini, i sistemi dinamici prevedibili non sono molto sensibili agli errori di misurazione.

L'operazione di stiramento e piegatura di un attrattore caotico elimina sistematicamente l'informazione iniziale e la sostituisce con informazione nuova: lo stiramento amplifica le indeterminazioni su piccola scala, la piegatura avvicina traiettorie molto lontane tra loro e cancella l'informazione su grande scala. Quindi gli attrattori caotici si comportano come una sorta di pompa, poiché portano a manifestazione macroscopica le fluttuazioni microscopiche. È chiaro allora che non può esistere alcuna soluzione esatta, alcuna scorciatoia per prevedere il futuro. Dopo un breve intervallo di tempo l'indeterminazione corrispondente alla misura iniziale ricopre tutto l'attrattore e tutta la capacità di previsione è perduta: non vi è più alcun legame causale tra passato e futuro.

Gli attrattori caotici sono frattali, cioè oggetti che rivelano un numero sempre più grande di particolari via via che vengono ingranditi. Il caos genera frattali in modo naturale. Le traiettorie vicine si allontanano, ma perché il moto resti finito prima o poi devono ripiegarsi e riavvicinarsi tra loro. Ciò avviene ripetutamente e si generano così pieghe dentro le pieghe, all'infinito. Michel Hénon, dell'Osservatorio di Nizza, ha scoperto una regola semplice che stira e piega il piano, portando ogni punto in una posizione nuova. Da un unico punto iniziale sono stati tracciati tutti i punti successivi ottenuti applicando ripetutamente la regola di Hénon. La figura geometrica risultante (*a*) è un esempio di attrattore caotico. In *b* il quadratino rosso è ingrandito dieci volte. La ripetizione del procedimento (*c*, *d*) rivela nei particolari la struttura microscopica dell'attrattore. L'illustrazione in basso raffigura un bacino di attrazione della mappa di Hénon.

Gli attrattori caotici fungono localmente da amplificatori di rumore. Una piccola fluttuazione, dovuta magari a rumore termico, provoca dopo breve tempo un ampio scostamento nella posizione dell'orbita. Ma gli attrattori caotici differiscono dagli ordinari amplificatori di rumore per un aspetto importante. Poiché si suppone che lo stiramento e la piegatura siano ripetuti e continui, qualsiasi minuscola fluttuazione finisce per dominare il moto e il comportamento qualitativo è indipendente dal livello di rumore. Perciò non è possibile «calmare» direttamente i sistemi caotici, per esempio abbassandone la temperatura. I sistemi caotici generano aleatorietà di per sé, senza bisogno di alcun apporto aleatorio esterno. Il comportamento aleatorio scaturisce da qualche cosa di più di una semplice amplificazione degli errori e da una perdita della capacità di previsione; è dovuto alle orbite complesse generate da stiramento e piegatura.

Si osservi che tanto il comportamento caotico quanto quello non caotico si possono presentare in sistemi conservativi, cioè privi di dissipazione. Qui le orbite non tendono a disporsi su un attrattore, ma sono confinate su una superficie di energia. La dissipazione, però, è importante in molti sistemi reali, se non in tutti, ed è naturale aspettarsi che il concetto di attrattore abbia un'utilità generale.

Gli attrattori caotici con poche dimensioni hanno inaugurato un nuovo dominio della teoria dei sistemi dinamici, ma resta da stabilire se abbiano a che fare con l'aleatorietà dei sistemi fisici. La prima prova sperimentale a sostegno dell'ipotesi che alla base del moto aleatorio dei fluidi vi siano attrattori caotici fu piuttosto indiretta. L'esperimento fu condotto nel 1974 da Jerry P. Gollub dell'Haverford College e da Harry L. Swinney dell'Università del Texas ad Austin. La prova risultò indiretta perché i ricercatori si concentrarono non sull'attrattore in sé, ma sulle proprietà statistiche che lo caratterizzano.

Il sistema studiato era una cella di Couette, che è costituita da due cilindri concentrici; lo spazio fra i due cilindri è occupato da un fluido e uno o entrambi i cilindri vengono fatti ruotare a una certa velocità angolare. All'aumentare della velocità angolare il fluido manifesta configurazioni di flusso sempre più complesse, che dipendono dal tempo in modo complicato (*si veda l'illustrazione di questa pagina*). Gollub e Swinney in sostanza misuravano la velocità del fluido in un dato punto. All'aumentare della velocità di rotazione, osservarono prima una transizione da una velocità costante nel tempo a una velocità che variava periodicamente e infine a una velocità che variava in modo non periodico. Il cuore dell'esperimento era questa transizione al moto non periodico.

L'esperimento doveva distinguere tra due modelli teorici che prevedevano un diverso comportamento del fluido al variare della velocità di rotazione. Il modello di Landau per il moto aleatorio del fluido prevedeva che all'aumentare della velocità venisse eccitato un numero sempre maggiore di oscillazioni indipendenti del fluido. L'attrattore associato era un toro a molte dimensioni. Il modello di Landau era stato messo in discussione da David Ruelle dell'Institut des Hautes Études Scientifiques, presso Parigi, e da Floris Takens, dell'Università di Groningen, nei Paesi Bassi. Le loro argomentazioni matematiche inducevano a credere che non fosse plausibile che nel moto dei fluidi si presentasse l'attrattore associato con il modello di Landau. I risultati suggerivano invece che qualunque toro multidimensionale potesse dar luogo a un attrattore caotico, come aveva postulato per primo Lorenz.

Gollub e Swinney scoprirono che a bassa velocità di rotazione il flusso non variava nel tempo: l'attrattore soggiacente era un punto fisso. All'aumentare della velocità angolare l'acqua cominciava a oscillare con una sola frequenza indipendente e l'attrattore corrispondente era un ciclo limite (un'orbita periodica); quando poi la rotazione diveniva ancora più veloce, l'oscillazione assumeva due frequenze indipendenti, e l'attrattore corrispondente era un toro bidimensionale. La teoria di Landau prevedeva che, aumentando ulteriormente la velocità di rotazione, questa tendenza di comportamento venisse confermata, nel senso che dovessero via via comparire altre frequenze distinte. In corrispondenza di una velocità di rotazione critica, invece, compariva all'improvviso una gamma di frequenze continua. Quest'osservazione era in accordo con il «flusso non periodico deterministico» di Lorenz, e corroborava la sua idea che alla base della turbolenza dei fluidi vi siano gli attrattori caotici.

L'analisi di Gollub e Swinney avvalorava l'idea che gli attrattori caotici potessero essere alla base di alcuni moti aleatori dei fluidi, ma il loro lavoro non fu certo conclusivo. Sarebbe auspicabile dimostrare esplicitamente l'esistenza nei dati sperimentali di un attrattore caotico semplice. In genere un esperimento non registra tutti gli aspetti di un sistema, ma solo alcuni. Gollub e Swinney non riuscirono a registrare tutto il flusso della cella di Couette, ma solo la velocità del fluido in un dato punto. Il compito dello studioso è «ricostruire» l'attrattore dai dati limitati ottenuti. E ciò non è sempre possibile; se l'attrattore è troppo complicato qualcosa andrà perduto. In certi casi, però, è possibile ricostruire la dinamica a partire da dati limitati.

Una tecnica introdotta da noi e posta su solide basi matematiche da Takens consentì di ricostruire lo spazio degli stati e di cercare gli attrattori caotici. L'idea di base è che l'evoluzione di ciascuna componente di un sistema è determinata da quelle componenti con le quali interagisce. Pertanto nella storia di ciascuna componente sono contenute implicitamente informazioni sulle componenti con cui interagisce. Per ricostruire uno spazio degli stati «equivalente», si considera una singola componente e si trattano i valori misurati con ritardi temporali fissi (un secondo prima, due secondi prima, e così via, per esempio) come se fossero nuove dimensioni.

a

b

c

d

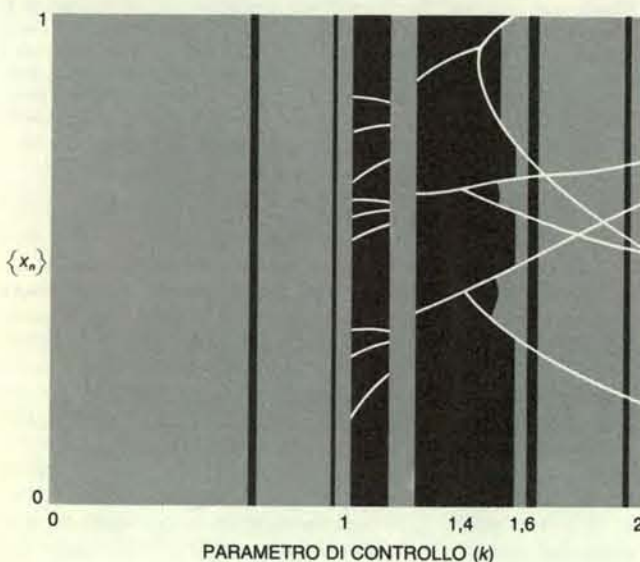
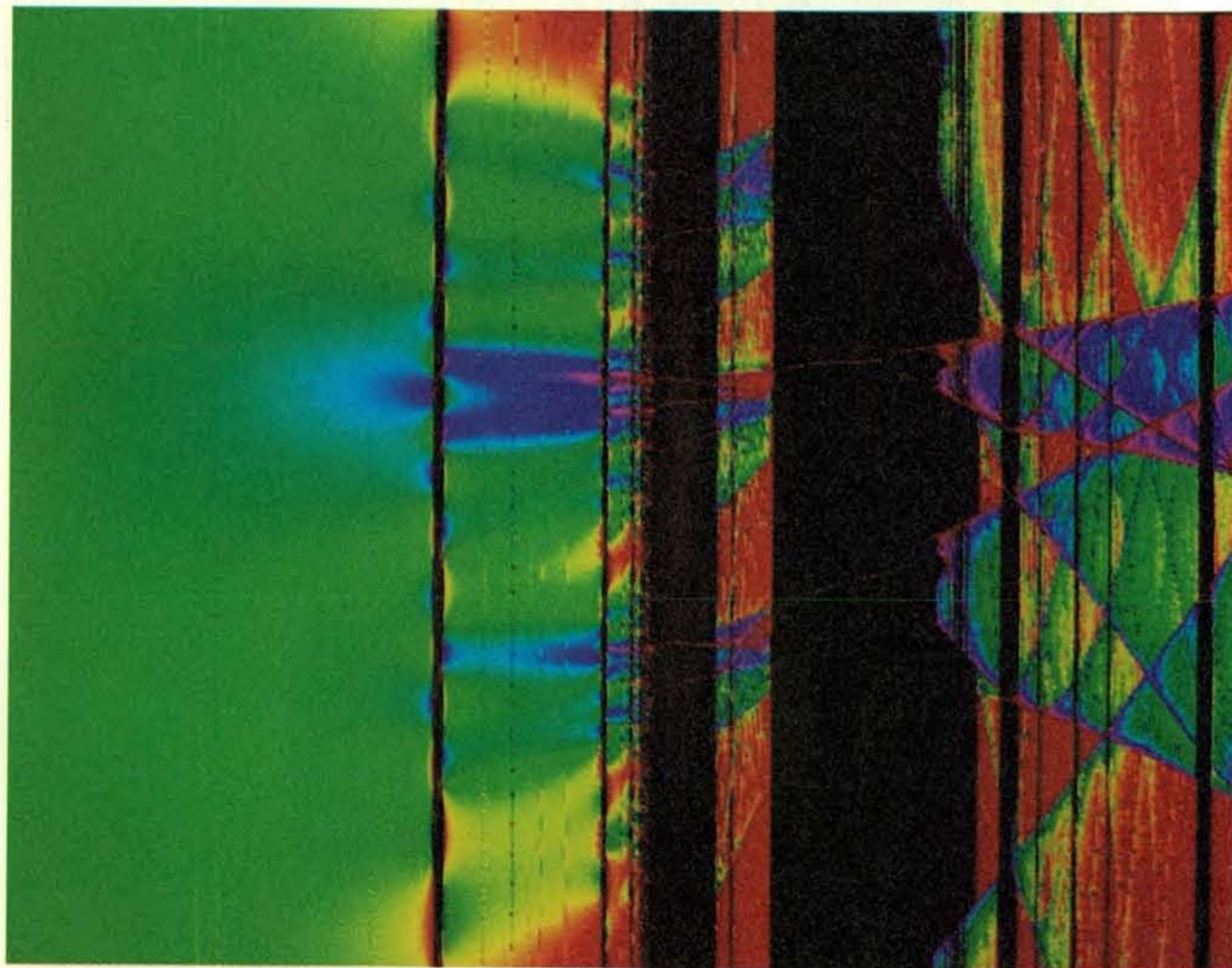
I risultati sperimentali confortano l'ipotesi che gli attrattori caotici siano alla base di alcuni tipi di moto aleatorio dei fluidi. L'acqua contenuta in una cella di Couette, che è costituita da due cilindri concentrici, è raffigurata in tre fotografie in successione. Lo spazio fra i due cilindri viene riempito d'acqua; quindi il cilindro interno viene posto in rotazione con una data velocità angolare (*a*). All'aumentare di questa velocità, il fluido manifesta un moto dall'andamento sempre più complesso (*b*): prima irregolare (*c*) e infine caotico (*d*).

Un rubinetto che gocciola è un esempio molto comune di sistema che può subire una transizione caotica. L'attrattore soggiacente viene ricostruito riportando a coppie in un grafico gli intervalli di tempo fra gocce successive, come appare nella parte in alto dell'illustrazione. Gli attrattori ricostruiti a partire da un reale rubinetto che gocciola (*a*, *c*) non sono di qualità inferiore agli attrattori generati mediante varianti della regola di Hénon (*b*, *d*). (L'attrattore di Hénon completo è a pagina 17.) Le illustrazioni *e* e *f* sono state ricostruite da flussi d'acqua ad alta velocità e rappresentano presumibilmente sezioni di attrattori caotici finora mai osservati. In ciascuno dei grafici sono state impiegate le coordinate ritardo temporale. La coordinata orizzontale è t_n, che è l'intervallo fra la goccia n e la goccia $n - 1$. La coordinata verticale è l'intervallo di tempo successivo, t_{n+1}, e la terza coordinata, ortogonale alla pagina, è t_{n+2}. Quindi ciascun punto è individuato da una terna di numeri (t_n, t_{n+1}, t_{n+2}) che sono stati tracciati per un insieme di 4094 campioni. Alle illustrazioni *b* e *d* è stato aggiunto rumore simulato.

I valori ritardati possono essere considerati come nuove coordinate, che definiscono un singolo punto in uno spazio degli stati multidimensionale. Ripetendo il procedimento e scegliendo i ritardi con riferimento a istanti diversi, si possono generare molti di questi punti. Si possono poi sfruttare altre tecniche per verificare se questi punti giacciono su un attrattore caotico o no. Benché questa rappresentazione sia per molti aspetti arbitraria, risulta che in essa vengono conservate le proprietà importanti di un attrattore, le quali non dipendono dai particolari della ricostruzione.

L'esempio che useremo per illustrare questa tecnica ha il pregio di essere noto e accessibile quasi a tutti. Molti sanno che le gocce che cadono da un rubinetto che gocciola presentano un andamento periodico. L'intervallo fra due gocce successive può essere regolarissimo, e molti, nelle notti d'insonnia, sono stati con l'orecchio teso ad aspettare che cadesse la goccia successiva. Meno noto è il comportamento di un rubinetto quando la velocità del flusso è un po' più elevata. Si può spesso individuare un regime in cui le gocce, pur cadendo sempre separate, cadono con un ritmo che non si ripete mai, come un tamburo che improvvisi cadenze sempre nuove. (È un esperimento che chiunque può effettuare facilmente; i rubinetti senza frangigetto funzionano meglio.) Il passaggio dal-

l'andamento ritmico a quello apparentemente aleatorio ricorda la transizione dal regime laminare a quello turbolento di un fluido. È possibile che alla base di questa aleatorietà vi sia un attrattore caotico semplice?

Lo studio sperimentale di un rubinetto che gocciola fu eseguito presso l'Università della California a Santa Cruz da uno degli autori (Shaw) in collaborazione con Peter L. Scott, Stephen C. Pope e Philip J. Martein. La prima variante dell'esperimento consisté nel far cadere le gocce da un normale rubinetto su un microfono e nel registrare i corrispondenti impulsi sonori. A pagina 19 sono illustrati alcuni risultati tipici di un esperimento un po' più raffinato. Un grafico in cui gli intervalli fra le gocce siano riportati a coppie è uno strumento efficace per ricavare una sezione dell'attrattore

La transizione al caos è illustrata da un diagramma di biforcazione, in cui una famiglia di attrattori (*asse verticale*) è rappresentata in funzione di un parametro di controllo (*asse orizzontale*). Il diagramma è stato generato da un semplice sistema dinamico che associa un numero a un altro. Il sistema dinamico impiegato è l'applicazione circolare ed è definito dall'equazione ricorsiva $x_{n+1} = \omega + x_n + k/2\pi \sin(2\pi x_n)$. Per ciascun valore scelto del parametro di controllo k un calcolatore ha tracciato l'attrattore corrispondente. I colori esprimono la probabilità di trovare i punti sugli attrattori: il rosso corrisponde a regioni visitate spesso, il verde a regioni visitate meno spesso e il blu a regioni visitate di rado. Al crescere di k da 0 a 2 (*si veda il disegno a sinistra*), il diagramma illustra due cammini verso il caos: una strada quasiperiodica (da $k = 0$ a $k = 1$, che corrisponde alla regione verde qui sopra) e una strada di «raddoppio del periodo» (da $k = 1,4$ a $k = 2$). La strada quasiperiodica è equivalente, sotto il profilo matematico, a un cammino che passa attraverso un attrattore a forma di toro. Nella strada del raddoppio del periodo, che è basata sull'attrattore a ciclo limite, le diramazioni si presentano a coppie, secondo la serie geometrica 2, 4, 8, 16, 32 e così via. Le iterazioni oscillano tra le coppie dei rami. (Per un valore particolare di k - per esempio 1,4 - le iterazioni visitano soltanto tre valori. Per k più grande, questa «orbita di periodo tre» raddoppia il suo periodo e visita sei valori; poi, con un ulteriore raddoppio ne visita 12, e via di questo passo.) Da ultimo la struttura delle diramazioni diventa così fine che ne emerge una struttura a banda continua: si raggiunge una soglia oltre la quale fa la sua comparsa il caos.

soggiacente. Nel regime periodico, per esempio, il menisco dal quale si staccano le gocce si sposta in un modo continuo e ripartitivo che potrebbe essere rappresentato da un ciclo limite nello spazio degli stati. Ma nell'esperimento reale questo movimento continuo è inaccessibile; vengono registrati solo gli intervalli tra il distacco di una goccia e della successiva. È come quando si applica una luce stroboscopica a un moto regolare su una curva chiusa: se la cadenza è scelta bene, si vede solo un punto fisso.

Il risultato più interessante dell'esperimento fu che nel regime non periodico del rubinetto gocciolante furono effettivamente trovati attrattori caotici. L'aleatorietà delle gocce sarebbe potuta derivare da cause invisibili, per esempio da piccole vibrazioni o da correnti d'aria. Ma, se fosse stato così, non vi sarebbe stata alcuna relazione particolare fra un intervallo e il successivo e il grafico dei dati presi a coppie avrebbe avuto l'aspetto di una chiazza informe. Il fatto che nei grafici compaiano delle strutture dimostra che l'aleatorietà possiede una base deterministica. In particolare molti gruppi di dati hanno una forma a ferro di cavallo che è caratteristica del semplice processo di stiramento e piegatura descritto in precedenza. Questa forma caratteristica può essere considerata come l'«istantanea» di una piegatura in atto, per esempio una sezione parziale dell'attrattore di Rössler (*si veda l'illustrazione a pagina 15*). Altri gruppi di dati appaiono più complicati; potrebbero essere sezioni di attrattori a più dimensioni. Attualmente la geometria degli attrattori a più di tre dimensioni è quasi completamente sconosciuta.

Se un sistema è caotico, quanto è caotico? Una misura del caos è l'«entropia» del moto, che è grosso modo il tasso medio di stiramento e piegatura, ovvero il tasso medio al quale viene generata l'informazione. Un'altra stima è la «dimensione» dell'attrattore. Se un sistema è semplice, il suo comportamento dovrebbe essere descritto da un attrattore dotato di poche dimensioni nello spazio degli stati, come negli esempi forniti in questo articolo. Per individuare lo stato di un sistema più complicato potrebbe essere necessario usare parecchi numeri, e l'attrattore corrispondente avrebbe un numero maggiore di dimensioni.

La tecnica di ricostruzione, affiancata da misurazioni di entropia e di dimensione, consente di riesaminare il moto del fluido studiato da Gollub e Swinney. Ciò è stato fatto da alcuni ricercatori del gruppo di Swinney in collaborazione con due degli autori (Crutchfield e Farmer). La tecnica di ricostruzione ci ha permesso di ottenere immagini dell'attrattore soggiacente. Queste immagini non ci danno la sensazionale dimostrazione di un attrattore di poche dimensioni, come gli studi di altri sistemi, per esempio il rubinetto gocciolante. Tuttavia, le misurazioni dell'entropia e della dimensione rivelano che il moto irregolare del fluido in prossimità della transizione nella cella di Couette può essere descritto da attrattori caotici. All'aumentare della velocità di rotazione della cella, aumentano anche l'entropia e la dimensione degli attrattori soggiacenti.

Negli ultimi anni si è dimostrato che un numero crescente di sistemi ha un'aleatorietà dovuta a un attrattore caotico semplice. Fra questi il moto di convezione di un fluido riscaldato in una piccola scatola, i livelli oscillanti della concentrazione in una reazione chimica per rimescolamento, la contrazione delle cellule cardiache di pollo e un gran numero di oscillatori elettrici e meccanici. Inoltre si è dimostrato che i modelli al calcolatore di certi fenomeni, che vanno dalle epidemie all'attività elettrica di una cellula nervosa alle oscillazioni delle stelle, possiedono questo tipo semplice di aleatorietà. Sono in corso anche esperimenti diretti a cercare il caos in aree disparate quali le onde cerebrali e l'economia.

Si deve tuttavia sottolineare che la teoria del caos è ben lungi dall'essere una panacea. Molti gradi di libertà possono anche dar luogo a moti complicati che sono effettivamente aleatori. Anche quando si sappia che un dato sistema è caotico, questo fatto da solo non dice molto. Un buon esempio è costituito dagli urti reciproci e dai rimbalzi delle molecole di un gas. Sapere che questo sistema è caotico non rende più facile la previsione del suo comportamento. Le molecole del sistema sono così numerose che si può solo sperare in una descrizione statistica e le proprietà statistiche fondamentali possono essere ricavate senza prendere in considerazione il caos.

Vi sono altri problemi non ancora studiati, nei quali il ruolo del caos è ignoto. Che cosa si può dire per esempio di configurazioni spaziali continuamente mutevoli, come le dune del Sahara, o della turbolenza pienamente sviluppata? Non è chiaro se configurazioni spaziali complesse possano essere descritte utilmente da un unico attrattore in un unico spazio degli stati. Comunque l'esperienza con gli attrattori più semplici può forse contribuire a guidarci verso un quadro più articolato, che può comprendere interi complessi di forme deterministiche mobili nello spazio e affini agli attrattori caotici.

L'esistenza del caos ha conseguenze per il metodo scientifico stesso. Il metodo classico per verificare una teoria è fare previsioni e confrontarle poi con i dati sperimentali. Ma se i fenomeni sono caotici, le previsioni a lungo termine sono intrinsecamente impossibili e nel giudicare la bontà della teoria si deve tener conto anche di questo. Quindi il procedimento di verifica di una teoria diviene un'operazione molto più delicata, basata su proprietà statistiche e geometriche anziché su previsioni particolareggiate.

Il caos rappresenta una nuova sfida per il punto di vista riduzionistico, secondo cui un sistema può essere compreso scomponendolo e poi studiandone le singole parti. Questo punto di vista è stato largamente seguito nella scienza anche perché sono moltissimi i sistemi per i quali il comportamento del tutto è effettivamente la somma delle parti. Il caos, tuttavia, dimostra che un sistema può manifestare un comportamento complicato come risultato di un'interazione non lineare semplice tra poche componenti soltanto.

Il problema si sta facendo acuto in un'ampia gamma di discipline scientifiche, dalla descrizione della fisica microscopica alla costruzione di modelli per il comportamento macroscopico degli organismi biologici. Negli ultimi anni la capacità di ricavare conoscenze particolareggiate sulla struttura di un sistema ha compiuto progressi formidabili, ma la capacità di integrare queste conoscenze è stata ostacolata dall'assenza di una cornice concettuale adatta entro la quale descrivere il comportamento qualitativo. Anche quando si possegga una mappa completa del sistema nervoso di un organismo semplice, come il nematode studiato da Sidney Brenner dell'Università di Cambridge, non è possibile ricavarne il comportamento dell'organismo. Analogamente, la speranza che la fisica possa raggiungere la compiutezza grazie a una comprensione sempre più particolareggiata delle forze fisiche e dei costituenti fondamentali è infondata. L'interazione delle componenti a una data scala può provocare su scala più vasta un comportamento globale complesso che in generale non può essere ricavato dalla conoscenza delle singole componenti.

Spesso il caos è visto in termini delle limitazioni che comporta, come la mancanza di prevedibilità. Ma accade che la natura sfrutti il caos in modo costruttivo. Grazie all'amplificazione delle piccole fluttuazioni, esso può consentire ai sistemi naturali di accedere alla novità. Una preda che sfugga all'attacco di un predatore potrebbe usare un controllo caotico del volo come un elemento di sorpresa per sfuggire alla cattura. L'evoluzione biologica richiede la variabilità genetica, e il caos offre un mezzo per la strutturazione delle variazioni aleatorie, fornendo così la possibilità di porre la variabilità sotto il controllo dell'evoluzione.

Anche il progresso intellettuale è basato sull'introduzione di idee nuove e su nuove connessioni fra idee vecchie. La creatività innata potrebbe essere basata su un processo caotico, che amplifica selettivamente piccole fluttuazioni e le foggia in stati mentali macroscopici coerenti che vengono esperiti come pensieri. In certi casi i pensieri possono essere decisioni o essere percepiti come un esercizio della volontà. Sotto questa luce, il caos fornisce un meccanismo che spiega il libero arbitrio nell'ambito di un mondo retto da leggi deterministiche.

Determinismo e caos

di Vincent Croquette

L'evoluzione di certi sistemi fisici, per esempio un pendolo, un circuito elettrico semplice, una sfera in caduta libera, può essere descritta con l'aiuto di una grandezza fisica misurabile (la variabile), del suo tasso di variazione (la derivata di questa variabile in rapporto al tempo), ed eventualmente del tasso di variazione del tasso di variazione della variabile (la sua derivata seconda in rapporto al tempo). La soluzione dell'equazione differenziale che lega queste diverse grandezze consente di descrivere perfettamente l'evoluzione successiva del sistema, che viene allora detto deterministico. Nel caso della sfera in caduta libera, l'equazione differenziale del moto è particolarmente semplice, in quanto traduce il fatto che l'accelerazione del corpo (la derivata seconda della sua posizione in rapporto al tempo) è costante e uguale al peso dell'unità di massa. Quindi, se si conoscono la posizione e la velocità della sfera a un istante dato (le condizioni iniziali), si può inserirle nella soluzione generale dell'equazione del moto per determinare la posizione e la velocità della sfera a qualsiasi istante successivo.

Ancora fino a poco tempo fa, i fisici ritenevano che i sistemi deterministici retti da un'equazione differenziale perfettamente nota avessero soluzioni regolari, come nel caso delle oscillazioni di un pendolo o della traiettoria di una sfera in caduta libera. Osservazioni successive, però, hanno dimostrato che quando l'equazione differenziale del sistema è «non lineare» il sistema stesso può avere comportamenti caotici: a prima vista, si potrebbe pensare che sia allora il caso a dirigere il suo moto. Una «equazione differenziale non lineare» è un'equazione in cui compaiono termini che non sono più semplicemente proporzionali alla variabile o a una delle sue derivate. L'equazione differenziale della sfera in caduta libera è effettivamente lineare, ma quella del pendolo semplice non lo è perché la forza che tende a riportare il pendolo nella sua posizione d'equilibrio varia al variare del seno dell'angolo formato dal pendolo rispetto alla verticale. Se per angoli piccoli si può sostituire la funzione seno con una semplice legge lineare, per angoli ampi questa approssimazione diventa molto scadente.

È occorso molto tempo perché si imponesse l'idea che i sistemi deterministici non sono retti da equazioni con soluzioni regolari, conoscibili con precisione infinita, in qualche modo perfette, benché non mancassero i controesempi, il più famoso dei quali è sicuramente la turbolenza nei fluidi. Essa presenta però una proprietà particolare che permette di spiegare il carattere caotico delle soluzioni: in un fluido un'infinità di configurazioni, o modi, possono diventare instabili e il fisico sovietico Lev Landau, intorno al 1950, aveva descritto la turbolenza come la manifestazione simultanea dell'insieme di questi modi, ciascuno associato a una frequenza particolare. Nel 1963 Edward Lorenz diede un importante contributo al progresso delle conoscenze sull'argomento: per definire un modello dei moti convettivi di uno strato d'aria propose un sistema di equazioni molto semplici e perfettamente deterministiche le cui soluzioni, caotiche, descrivevano la turbolenza dello strato d'aria senza far intervenire questa infinita serie di modi. Da allora si è voluto vedere il caos un po' dappertutto nei sistemi deterministici semplici, e oggi i fisici si interessano a numerosi fenomeni che, per quanto individuati, erano stati trascurati perché se ne attribuiva il carattere caotico all'inadeguatezza delle condizioni sperimentali e non al carattere non lineare della loro equazione.

La semplicità nel caos

L'entusiasmo dei fisici per i sistemi stocastici è legato anche al fascino delle strutture che questi sistemi possono avere: il carattere deterministico che si riteneva spesso cancellato dalla stocasticità si manifesta, in certo qual modo, nell'ordine relativo che traspare dalle soluzioni caotiche. Il nome di «attrattori strani» dato alle rappresentazioni matematiche di queste soluzioni traduce bene quel fascino. Descriveremo più avanti alcuni di questi attrattori.

Un semplice esempio matematico che illustra bene questi diversi aspetti della stocasticità è quello, ormai classico, delle iterazioni di una funzione la cui rappresentazione grafica ha la forma di una campana. Le funzioni a una variabile sono espressioni matematiche il cui valore è determinato dal valore attribuito a una sola grandezza. Data quindi una funzione f della sola variabile x, il processo di iterazione, a partire da un valore iniziale x_0, consiste nel calcolare $f(x_0)$ - il valore assunto dalla funzione f quando la variabile x ha valore x_0 - e nel chiamare x_1 questo valore di $f(x_0)$; il valore del terzo termine x_2 di questa successione è allora uguale a $f(x_1)$, e così via. Per certi tipi di funzioni f dipendenti da un parametro λ, come per esempio la funzione $f(x) = 4\lambda x(1-x)$, la successione x_n dei valori consecutivi della funzione diventa caotica al di là di un certo valore di λ; questa fase caotica è preceduta da una fase regolare e la transizione verso il caos avviene in modo molto particolare: nel caso della funzione $f(x) = 4\lambda x(1-x)$ la successione dei valori x_n, per un valore iniziale x_0 compreso tra zero e uno, tende verso un unico valore limite quando λ è inferiore a 3/4. Per valori di λ leggermente superiori a 3/4, la successione non ammette più un limite unico, ma i suoi termini si avvicinano in alternativa a due valori: è avvenuto quello che si chiama un primo sdoppiamento. Quando il parametro λ diventa superiore a 0,086237, la successione oscilla intorno a quattro valori: è comparso un secondo sdoppiamento. Gli sdoppiamenti si succedono poi a «cascata» all'aumentare del parametro λ: un terzo sdoppiamento dà luogo a un ciclo di otto valori e così via, finché compare lo stato caotico per un valore di λ corrispondente a un'infinità di sdoppiamenti.

Questo esempio può sembrare troppo semplice e lontano dai sistemi fisici reali.

Eppure, Albert Libchaber e Jean Maurer dell'École Normale Supérieure hanno osservato per primi, nel corso di esperimenti di turbolenza nell'elio e nel mercurio, una situazione del tutto analoga di transizione al caos. In seguito altri fisici hanno rilevato questo tipo di fenomeno nei fluidi più comuni, per esempio l'acqua (Jerry P. Gollub e Marzio Giglio).

Un dispositivo estremamente semplice, costituito da un'ago di bussola posto tra due campi magnetici, permette di presentare i principali aspetti dei sistemi deterministici in cui compare la stocasticità (*si veda l'illustrazione in questa pagina*). Il fatto che questo esempio si ricolleghi alla meccanica classica è lungi dall'essere fortuito: anche se viene ricordato raramente, la meccanica classica ha svolto un ruolo fondamentale nella comprensione di questi sistemi; il presente articolo vuole porre rimedio a questo scarso riconoscimento.

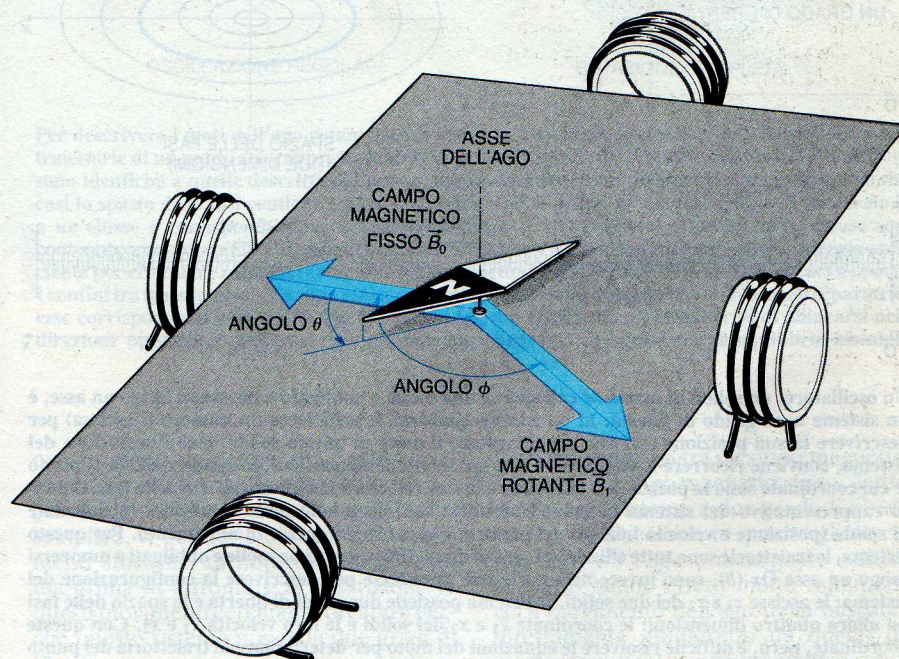
L'eclissi della meccanica classica

I successi conseguiti verso il 1930 dalla meccanica quantistica nella descrizione dei fenomeni microscopici hanno relegato in secondo piano la meccanica classica. Tuttavia, numerosi enigmi della meccanica classica, direttamente legati alla stocasticità, conservano la loro importanza. Non è stato risolto, per esempio, l'importante problema dei tre corpi, nei confronti del quale Henri Poincaré manifesta, in una memoria del 1889, un profondo disagio: «Quale sarà il movimento di n punti materiali che si attirino reciprocamente in ragione diretta della loro massa e in ragione inversa del quadrato delle distanze? Se $n = 2$, cioè se si ha a che fare con un pianeta isolato e con il Sole, trascurando le perturbazioni dovute agli altri pianeti, l'integrazione è facile: i due corpi descrivono delle ellissi, conformandosi alle leggi di Keplero. La difficoltà comincia se il numero n dei corpi è uguale a tre; il problema dei tre corpi ha sfidato finora tutti i tentativi di analisi, dato che l'integrazione rigorosa e completa è manifestamente impossibile.» Poincaré propone così il concetto di non integrabilità di un problema di meccanica, cioè il fatto che non si potrà mai trovare una soluzione esatta a questo problema. L'esempio della bussola posta tra due campi magnetici precisa il concetto di problema integrabile e mostra perché certi problemi, analoghi a quello dei tre corpi, non siano integrabili.

I problemi integrabili

Per risolvere un problema di meccanica si cerca innanzitutto il suo numero di gradi di libertà, cioè il numero di variabili di cui si ha bisogno per descrivere la configurazione del sistema considerato a un istante dato. Un punto materiale che si sposti lungo un asse possiede un solo grado di libertà: basta conoscere la sua ascissa in rapporto a un punto scelto come origine per determinare senza alcuna ambiguità la sua posizione; un pianeta considerato come un punto materiale possiede tre gradi di libertà; un sistema di due corpi, sei gradi di libertà; uno di tre corpi, nove gradi di libertà. La configurazione di un sistema a N gradi di libertà è dunque definita da N variabili dette di posizione, ma per determinare lo stato fisico del sistema occorre inoltre conoscere il valore delle velocità associate ai diversi gradi di libertà o, più esattamente, la quantità di moto di ciascuno dei corpi (cioè la massa del corpo moltiplicata per la sua velocità). Lo stato del sistema è dunque definito da $2N$ coordinate, di cui N sono coordinate di posizione e N coordinate di quantità di moto. Per descrivere l'evoluzione del sistema conviene rappresentare il suo stato come un punto in uno spazio a $2N$ dimensioni, con la posizione del punto definita dalle $2N$ coordinate del sistema. Lo spazio così definito si chiama spazio delle fasi e la traiettoria del punto che rappresenta il sistema in questo spazio è determinata, da una parte, dalle equazioni del moto, che formano un sistema di $2N$ equazioni differenziali che collegano le $2N$ coordinate del sistema (in cui sono compresi i diversi accoppiamenti tra i gradi di libertà), ma anche, d'altra parte, dal punto di partenza di questa traiettoria, ossia dalle condizioni iniziali del sistema.

Integrare il problema consiste nel risolvere il sistema di $2N$ equazioni differenziali che descrive la traiettoria nello spazio delle fasi. La difficoltà della soluzione deriva dagli accoppiamenti tra le diverse equazioni; il metodo per risolvere questo sistema di $2N$ equazioni consiste nel trovare una trasformazione di coordinate che permetta al sistema di $2N$ equazioni di scindersi in N sistemi di due equazioni equivalenti a quelle di un sistema a un grado di libertà. Consideriamo, per esempio, il caso di due oscillatori armonici identici formati da due masse mobili lungo un asse, ciascuna delle quali collegata a una molla (*si veda l'illustrazione a pagina 36*). Se questi due oscillatori, che formano un sistema a due gradi di libertà, sono accoppiati tramite una terza molla, la posizione x_1, della prima massa dipenderà dalla sua velocità v_1 ma anche dalla posizione x_2 e dalla velocità v_2 della seconda massa. Se ora scriviamo le equazioni del moto nel sistema di coordinate $X_1 = x_1 + x_2$ e $X_2 = x_1 - x_2$, $V_1 = v_1 + v_2$ e $V_2 = v_1 - v_2$, le equazioni che contengono X_1 e V_1 sono separate da quelle che contengono X_2 e V_2: abbiamo trovato, tramite queste trasformazioni, i modi caratteristici del sistema, cioè due nuovi oscillatori indipendenti.

In termini generali, un sistema a N gradi di libertà è integrabile se lo si può scomporre in N sistemi indipendenti a un grado di libertà. Il problema è trovare le coordinate di questi modi caratteristici. La cosa è sempre possibile quando il sistema è lineare, e quindi tutti i problemi lineari sono integrabili. Nel caso dei sistemi non lineari invece questa determinazione dei modi caratteristici è molto più difficile e,

L'ago di una bussola posto all'interno di due campi magnetici, uno fisso B_0 e l'altro rotante B_1, può avere, quando lo si avvia fissando le condizioni iniziali, un comportamento caotico. In questo tipo di comportamento, i moti dell'ago sembrano aleatori (o caotici); eppure, l'equazione del moto dell'ago è perfettamente nota (o deterministica). Due coppie di bobine di Helmholtz (le cui dimensioni sono state qui ridotte) producono, se sono percorse da correnti elettriche sinusoidali, il campo magnetico fisso B_0 e il campo magnetico rotante B_1. Questi due campi magnetici si trovano nel piano di rotazione dell'ago, la cui posizione è data dall'angolo θ tra esso e il campo magnetico fisso B_0. Quando il campo magnetico B_1 ruota a velocità angolare ω_0 costante, l'angolo φ tra i due campi magnetici è uguale al prodotto della velocità angolare ω_0 per il tempo. Questo semplice dispositivo permette di osservare i moti caotici di un sistema che è tuttavia perfettamente deterministico.

nella grande maggioranza dei casi, impossibile. La nozione di non integrabilità è dunque strettamente legata al carattere non lineare del sistema, per quanto si conoscano sistemi non lineari che sono integrabili, come per esempio il pendolo semplice.

Il fatto che sia possibile scomporre un sistema a N gradi di libertà in N sistemi separati significa che esistono N costanti del moto. Tutti i sistemi meccanici possiedono almeno una costante del moto, che generalmente è l'energia: ne risulta che la traiettoria del sistema nello spazio delle fasi deve appartenere a una ipersuperficie di energia costante a $2N-1$ dimensioni. Quando il sistema è integrabile esistono N costanti del moto e la traiettoria del punto rappresentativo deve appartenere a un'ipersuperficie molto più «piccola», di N dimensioni soltanto. Questa ipersuperficie è un po' particolare essendo il «prodotto» delle N traiettorie dei sistemi a un grado di libertà; vedremo che finché l'evoluzione di un sistema a un grado di libertà resta circoscritta, cioè confinata in un volume finito dello spazio delle fasi, la traiettoria del sistema appartiene alla famiglia della circonferenza, e l'ipersuperficie che rappresenta la sua evoluzione è allora il prodotto di N circonferenze, cioè un toro di dimensioni N (il toro comune, come per esempio una camera d'aria, è in realtà il prodotto di due circonferenze indipendenti).

Nel caso che il sistema non sia integrabile, è impossibile trovare N costanti del moto; di solito ne esiste una sola, vale a dire l'energia. La traiettoria è allora molto più «libera» in quanto appartiene a una ipersuperficie di dimensione $2N-1$. Ne deriva che, all'interno dei sistemi a N gradi di libertà (per i quali la dimensione dello spazio delle fasi è uguale a $2N$), la classe dei problemi integrabili appare come un sottoinsieme molto piccolo (se N è superiore a 1). Quando il sistema ha soltanto due gradi di libertà ($N = 2$) i problemi integrabili sono quelli le cui traiettorie nello spazio delle fasi si collocano su tori di dimensione due, mentre le traiettorie dei sistemi non integrabili a due gradi di libertà si collocano *a priori* su ipersuperfici a tre ($2 \times 2 - 1$) dimensioni. Pertanto, il numero di gradi di libertà è uguale o superiore a due e in genere i problemi di meccanica non sono più integrabili.

Un oscillatore, un solido di massa M collegato a una molla e obbligato a muoversi lungo un asse, è un sistema a un grado di libertà: in un istante qualsiasi, è sufficiente un numero (l'ascissa) per descrivere la sua posizione (*a*). Per rappresentare il moto di questo solido, cioè l'evoluzione del sistema, conviene ricorrere a un punto in uno spazio a due dimensioni, lo spazio delle fasi, punto le cui coordinate sono la posizione x del solido e la sua velocità $\dot{x}$. In questo spazio delle fasi, il punto rappresentativo del sistema descrive traiettorie che dipendono dalle condizioni iniziali date al solido (posizione e velocità iniziali) nel momento in cui lo si è messo in movimento. Per questo sistema, le traiettorie sono tutte ellissi. Nel caso di due solidi collegati da molle e obbligati a muoversi lungo un asse Ox (*b*), sono invece necessarie due grandezze per descrivere la configurazione del sistema: le ascisse x_1 e x_2 dei due solidi. Il sistema possiede due gradi di libertà e lo spazio delle fasi ha allora quattro dimensioni: le coordinate x_1 e x_2 dei solidi e le loro velocità $\dot{x}_1$ e $\dot{x}_2$. Con queste coordinate, però, è difficile risolvere le equazioni del moto per determinare la traiettoria del punto rappresentativo del sistema nello spazio delle fasi. Operando il cambiamento di coordinate $X = x_1 + x_2$ e $Y = x_1 - x_2$, si ottengono equazioni identiche a quella del caso precedente risolvibile: in questo nuovo sistema di coordinate, le soluzioni sono quelle di oscillatori indipendenti e le traiettorie si inscrivono su tori a due dimensioni dello spazio delle fasi a quattro dimensioni (qui non rappresentato). Allo stesso modo, i problemi integrabili a N gradi di libertà sono problemi per i quali si possono porre le equazioni del moto sotto forma di sistemi corrispondenti a N oscillatori indipendenti; le traiettorie, nello spazio delle fasi a $2N$ dimensioni, appartengono allora a tori di N dimensioni. (In realtà, il cambiamento di variabili è generalmente molto più difficile da trovare di quello del nostro esempio.) L'ago della bussola collocato nel solo campo magnetico fisso B_0 è un sistema integrabile; quando invece lo si colloca in due campi magnetici, uno fisso B_0 e uno rotante B_1, si ottiene un sistema a due gradi di libertà, non integrabile, in cui si manifesta il caos.

Le illusioni perdute dei deterministi

Che cosa significa che un sistema non è integrabile? Sono molti i casi in cui la matematica non consente la descrizione esatta dei fenomeni: non esiste, per esempio, un metodo per calcolare il valore esatto delle radici di un'equazione di n-esimo grado. Tutti noi conosciamo il metodo per trovare le radici dell'equazione di secondo grado $ax^2 + bx + c = 0$; ben pochi sono coloro che conoscono i metodi per risolvere un'equazione di terzo e quarto grado; quando poi il grado dell'equazione è superiore a quattro, non esiste più un metodo generale di risoluzione, benché sia dimostrato che le equazioni di n-esimo grado possiedono in generale n radici. In meccanica classica, il problema della non integrabilità è molto diverso: un sistema è non integrabile non soltanto perché ci manca un metodo di soluzione, ma soprattutto perché le sue traiettorie nello spazio delle fasi sono di tipo nuovo, sono traiettorie stocastiche. Ancora una volta fu Poincaré a osservare per primo questo tipo di soluzioni.

Per dimostrare come sia possibile che sistemi perfettamente deterministici abbiano soluzioni stocastiche e che cosa siano esattamente queste soluzioni, esaminiamo innanzitutto una delle proprietà essenziali delle soluzioni regolari: in esse si associano i movimenti di N oscillatori, ciascuno dei quali vibra alla propria frequenza. Quando si analizza il sistema nei termini di una delle grandezze dipendenti dal tempo, così come il nostro orecchio fa nel caso dei suoni, si ottiene la seguente corrispondenza: le soluzioni regolari sono come gli accordi in musica; l'energia è concentrata in piccolissimi intervalli di frequenza (le differenti frequenze degli oscillatori) e in intervalli corrispondenti alle frequenze delle loro armoniche (un'armonica è una vibrazione di frequenza pari a un multiplo intero della frequenza propria dell'oscillatore). Se si «ascoltasse» allo stesso modo un sistema stocastico, il nostro orecchio potrebbe forse distinguere note di altezza (frequenza) ben definita, ma sentiremmo anche un «rumore» abbastanza simile a quello di un torrente. Se si traccia in un grafico l'energia di vibrazione in funzione della frequenza si ottiene lo spettro di Fourier; per le soluzioni regolari lo spettro di Fourier è fatto di picchi infinitamente sottili, mentre gli spettri delle soluzioni stocastiche comportano anche righe larghe associate al rumore.

C'è un altro criterio che permette di distinguere le soluzioni regolari dalle soluzioni stocastiche: la sensibilità alle condizioni iniziali. Immaginiamo di costruire due bussole rigorosamente identiche e di porle in identiche configurazioni di campi magnetici, in condizioni iniziali rigorosamente uguali: i loro movimenti, regolari o stocastici che siano, saranno esattamente identici, il che ne evidenzia il carattere deterministico. Quando, caso più realistico, le condizioni iniziali dei due sistemi differiscono di pochissimo, l'evoluzione

relativa delle due bussole dipende in modo fondamentale dalla natura dei loro movimenti: se sono regolari, le bussole andranno progressivamente sfasandosi l'una rispetto all'altra, e la differenza tra i due movimenti crescerà linearmente nel tempo, come due orologi uno dei quali acceleri. In caso di traiettorie stocastiche, invece, la piccola differenza che esisteva all'inizio tra i due movimenti cresce esponenzialmente nel tempo, e i movimenti diventano diversissimi molto rapidamente, tanto che si stenta a credere che abbiano avuto qualcosa in comune. In meteorologia, Lorenz ha denominato «effetto farfalla» il fatto che l'evoluzione dell'atmosfera, di cui si intuisce il carattere caotico, venga radicalmente modificata da un cambiamento anche infimo delle condizioni iniziali, come potrebbe essere quello prodotto dal battito d'ali di una farfalla.

Attualmente i risultati ottenuti nell'ambito della stocasticità sono utilizzati in numerosi campi di studio: fisica dei plasmi, accelerazione di particelle, confinamento magnetico, stabilità delle fasce di Van Allen ecc. Inoltre, sono entrati a far parte di questo settore anche i sistemi dissipativi, cioè quelli in cui, contrariamente a quanto si studia in meccanica classica, l'energia non si conserva nel corso del tempo. Così, anche la turbolenza nei fluidi o nelle reazioni chimiche viene attualmente considerata da un'angolazione nuova e si può prevedere che nei prossimi anni numerosi ricercatori utilizzeranno questi risultati nei domini della fisica non lineare.

Per descrivere i moti dell'ago rotante senza attriti in un campo magnetico fisso B_0 si utilizzano le traiettorie di un punto che rappresenta il sistema in uno spazio detto spazio delle fasi. Le traiettorie sono identiche a quelle descritte dal punto rappresentativo di un pendolo semplice. In entrambi i casi lo spazio delle fasi contiene traiettorie di due tipi: le prime sono curve chiuse di forma simile a un'ellisse e corrispondenti a oscillazioni dell'ago (o del pendolo). Queste curve chiuse sono contenute in un dominio chiamato risonanza (*in colore*). Quando si fornisce all'ago energia sufficiente per effettuare rotazioni complete, la sua traiettoria nello spazio delle fasi è una curva aperta. I confini tra la risonanza e il dominio esterno contenente le traiettorie aperte si chiamano separatrici: esse corrispondono a moti dell'ago di energia appena sufficiente a permettergli di allinearsi nella direzione opposta a quella del campo magnetico fisso B_0, una posizione di equilibrio instabile.

Dal sistema integrabile al sistema non integrabile

Torniamo alla nostra bussola posta in un campo magnetico uniforme e stazionario, per esempio il campo terrestre: quali che siano le condizioni iniziali, l'ago si allinea con il campo magnetico dopo aver compiuto alcune oscillazioni da una parte e dall'altra. Se ruotasse senza alcun attrito, l'ago oscillerebbe indefinitamente perché costituisce un sistema del tutto equivalente al pendolo semplice: come quest'ultimo, l'ago della bussola possiede un grado di libertà e appartiene alla classe dei sistemi per i quali l'insieme dei casi integrabili ha la stessa dimensione dell'insieme delle soluzioni (per $N = 1$, si ha $N = 2N - 1$). Tutti i sistemi a un grado di libertà sono integrabili e anche il pendolo semplice, malgrado il suo carattere non lineare, lo è; come il nostro ago soggetto a un campo magnetico fisso, il pendolo avrà solo movimenti regolari.

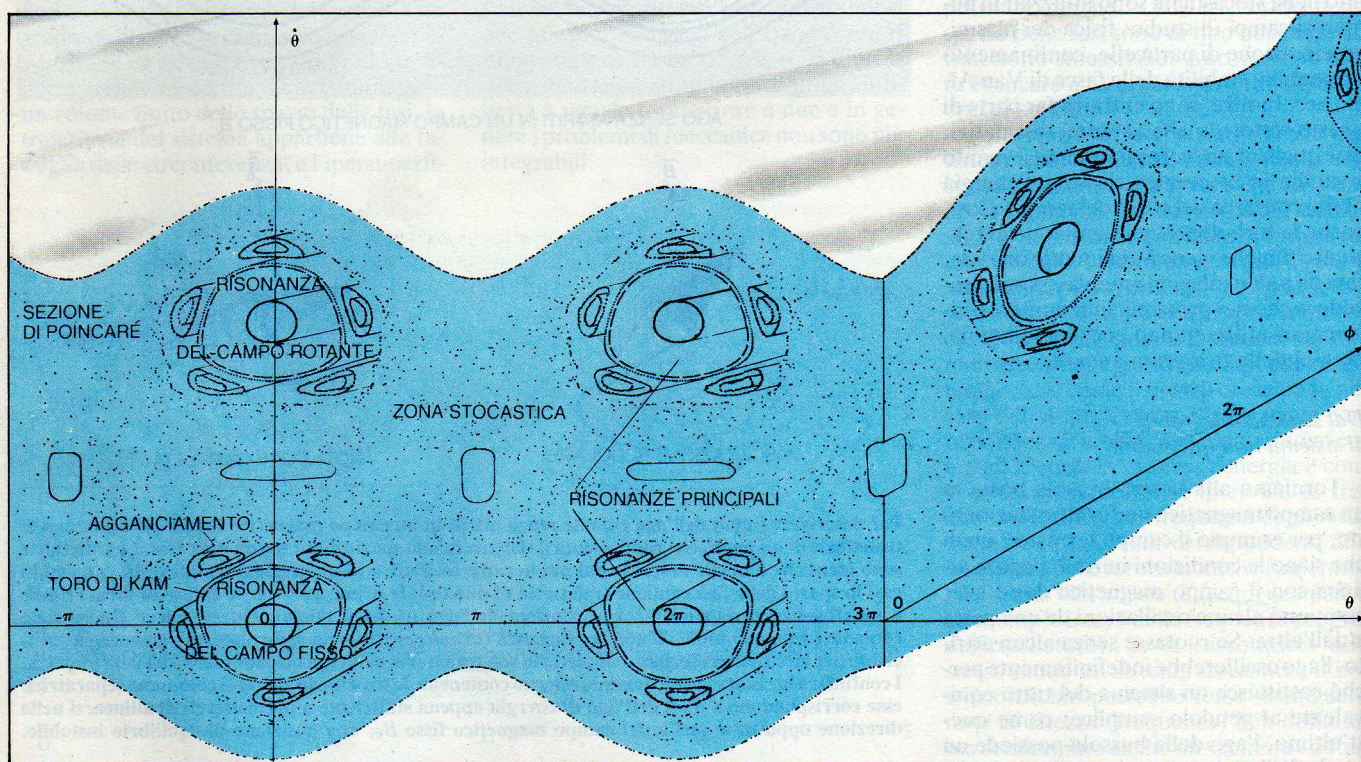
Per far comparire movimenti stocastici occorre aumentare il numero di gradi di libertà: a questo scopo, si aggiunge al campo magnetico fisso un campo magnetico rotante (*si veda l'illustrazione a pagina 35*). In pratica, per aggiungere un campo magnetico rotante al campo magnetico terrestre (fisso) che agisce sull'ago, basta mettere una barretta magnetica sul piatto di un giradischi il cui asse coincida con quello dell'ago della bussola; si produce in questo modo un campo magnetico orizzontale in rotazione alla stessa velocità del piatto. Il sistema così realizzato, molto semplice, ci permette di osservare movimenti stocastici di notevole complessità ma prima di affrontare la descrizione di questi movimenti cerchiamo di capire perché questo sistema possieda due gradi di libertà.

Per restare nell'ambito della meccanica classica, supporremo che il piatto del giradischi ruoti senza attriti dopo essere stato avviato con velocità angolare ω_0 e che quindi continui a ruotare a velocità costante grazie alla sua inerzia. In queste condizioni si può descrivere la configurazione del sistema con due variabili: l'angolo θ tra l'ago e il campo magnetico fisso, e l'angolo ϕ tra il campo magnetico rotante e il campo magnetico fisso. Tuttavia, questo sistema a due gradi di libertà è un po' particolare, in quanto gli accoppiamenti tra l'ago e il magnete solidale al piatto hanno effetti molto asimmetrici: il movimento dell'ago della bussola è direttamente influenzato da quello del piatto (tramite il campo magnetico rotante), mentre il movimento del piatto è influenzato in maniera ridottissima da quello dell'ago. Ciò è tanto più vero in quanto il piatto ha una massa di molto superiore all'ago e la sua inerzia è molto più grande: nel caso limite in cui la massa del piatto sia infinita, questo si comporta come un sistema isolato e il suo moto è solo una rotazione regolare; l'angolo ϕ è uguale al

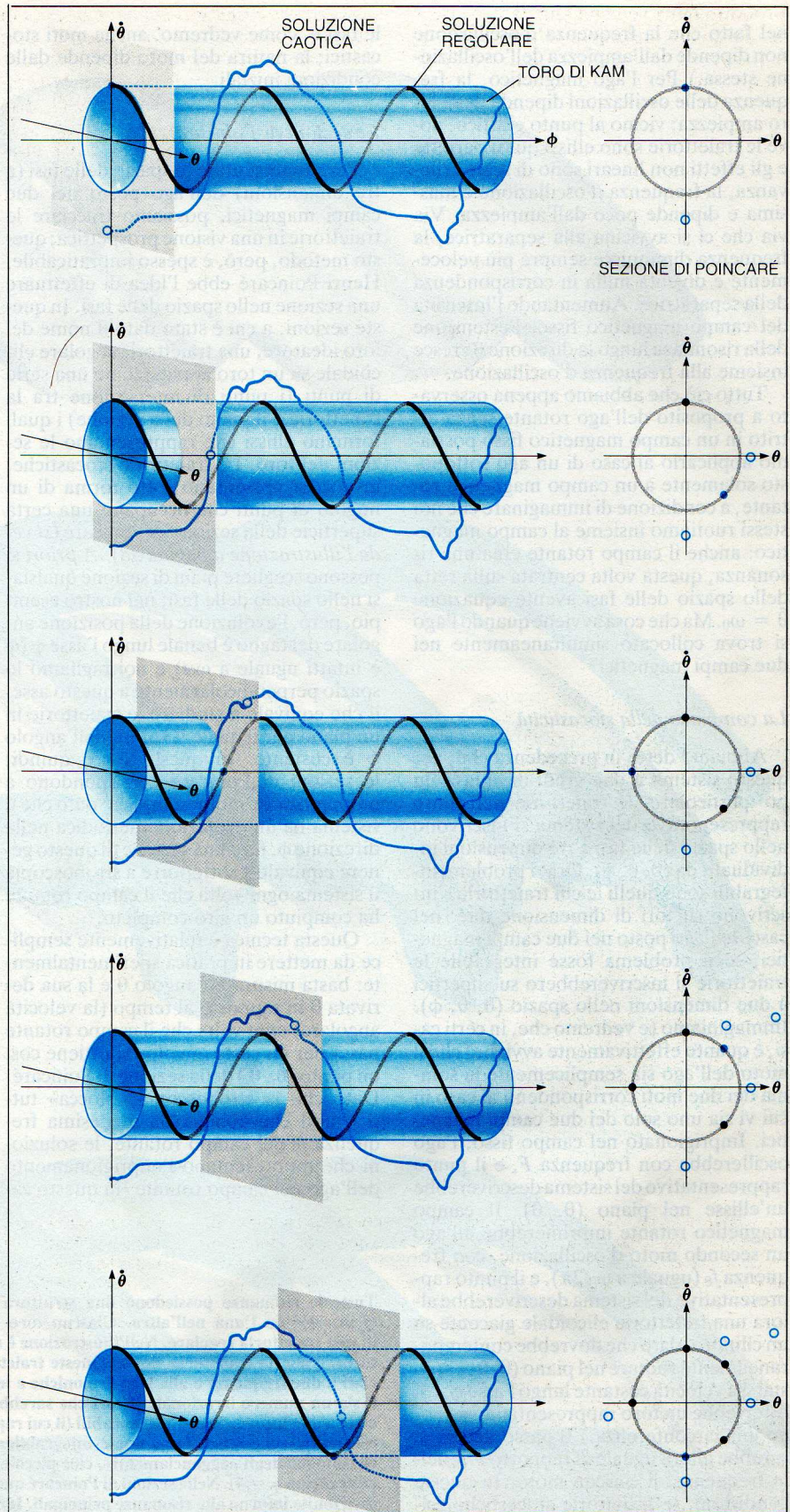
prodotto della velocità angolare ω_0 per il tempo t. Nello spazio delle fasi $(\theta, \dot{\theta}, \phi, \dot{\phi})$ a quattro dimensioni, l'evoluzione del sistema avviene in un sottospazio a tre dimensioni un po' particolare: dato che ϕ è uguale a $\omega_0 \times t$, la derivata $\dot{\phi}$ dell'angolo ϕ è uguale alla costante ω_0 e questo sottospazio è quindi $(\theta, \dot{\theta}, \phi)$, con $\dot{\theta}$ che indica la velocità angolare dell'ago della bussola. Inoltre, l'evoluzione del punto rappresentativo del sistema lungo l'asse ϕ è molto semplice perché l'angolo ϕ varia in modo lineare nel tempo.

Fino a questo punto abbiamo trascurato l'influenza degli attriti che dissipano energia nei sistemi fisici, ma è proprio l'esistenza di questi attriti a determinare la natura di un sistema. Un sistema è non dissipativo, o hamiltoniano, in assenza di attrito, come nel caso dei sistemi celesti; i sistemi a noi più accessibili, invece, sono di solito dissipativi perché presentano attriti: è il caso dell'ago della bussola. Noi prenderemo in esame in primo luogo le soluzioni del sistema non dissipativo ottenuto tramite simulazioni al calcolatore.

Il campo magnetico fisso

Consideriamo innanzitutto il caso in cui il campo magnetico rotante sia nullo; come abbiamo visto, il problema è allora analogo allo studio del movimento di un pendolo: lo spazio delle fasi possiede due dimensioni, una individuata dall'angolo θ tra l'ago e il campo magnetico fisso, l'altra dalla velocità angolare $\dot{\theta}$ dell'ago. Dato che due configurazioni che differiscano solo per un numero intero di giri dell'ago sono identiche, lo spazio delle fasi è periodico lungo l'asse θ, ossia la sua configurazione ricorre ogni volta che si aumenta θ di 2π. Ci limiteremo quindi a studiare questa configurazione per valori dell'angolo θ compresi tra $-\pi$ e $+\pi$. Grazie alle equazioni del moto, si ottengono le differenti traiettorie del punto rappresentativo del sistema nello spazio delle fasi (*si veda l'illustrazione a pagina 37*). Queste traiettorie sono molteplici in quanto, in assenza d'attrito, ne esiste una per ciascun valore dell'energia totale: a ciascuna coppia $(\theta_0, \dot{\theta}_0)$ di condizioni iniziali è associata un'energia e una traiet-

Lo spazio delle fasi del sistema formato dall'ago rotante senza attriti in un campo magnetico fisso B_0 e in un campo magnetico B_1 rotante a velocità costante ω_0 ha quattro dimensioni: l'angolo θ tra la direzione dell'ago e il campo magnetico fisso B_0, la velocità angolare $\dot{\theta}$ dell'ago, l'angolo ϕ tra il campo magnetico rotante B_1 e il campo magnetico fisso B_0 e la velocità angolare $\dot{\phi}$ del campo magnetico rotante B_1. Dato che il campo magnetico B_1 ruota a velocità angolare ω_0 costante, ci si può limitare allo studio dello spazio delle fasi a tre dimensioni $(\theta, \dot{\theta}, \phi)$. Inoltre, dato che le configurazioni del sistema che differiscono tra loro solo per un numero intero di giri dell'ago corrispondono allo stesso stato, lo spazio delle fasi è periodico (nell'illustrazione sono rappresentati solo due periodi secondo θ e uno secondo ϕ). Lo spazio delle fasi comporta traiettorie regolari che si avvolgono su tori, detti tori di KAM (dal nome dei matematici Kolmogorov-Arnold-Moser) e corrispondenti a moti regolari dell'ago (il numero delle frequenze che compongono questi moti è limitato); lo spazio delle fasi comporta inoltre traiettorie stocastiche corrispondenti a moti disordinati dell'ago e non avvolte su tori di KAM. I tori si raggruppano in due risonanze principali: la risonanza intorno al campo magnetico fisso (*in basso*), nella quale l'ago è «imprigionato» intorno a questo campo, e la risonanza intorno al campo magnetico rotante (*in alto*), dove l'ago oscilla intorno al campo magnetico rotante accompagnandolo nel suo moto di rotazione. I tori di KAM posti al centro delle risonanze principali corrispondono a moti dell'ago composti da due frequenze fondamentali, il cui rapporto Γ è un numero irrazionale. La non integrabilità del problema fa sì che gli altri tori, associati a rapporti tra frequenze che sono numeri razionali, cedano il posto a zone stocastiche e ad «agganciamenti», cioè a incastri di piccoli tori eccentrici rispetto alle risonanze principali e costituiti dall'avvolgimento a elica di traiettorie regolari. Per queste traiettorie regolari avvolte sui tori degli agganciamenti, i moti dell'ago comportano ancora due frequenze: la frequenza d'oscillazione intorno al «punto ellittico» situato al centro dell'agganciamento, e la più piccola delle due frequenze che avrebbero formato il rapporto razionale nel problema integrabile. Dato che questo rapporto F/f_0 è uguale a un numero razionale p/q, la frequenza più alta è un multiplo della più bassa, la quale dunque è semplicemente una delle sue armoniche. I tori che avrebbero dovuto corrispondere alle separatrici delle risonanze (il confine dei domini in cui l'ago è imprigionato, cioè costretto a oscillare intorno a uno dei campi magnetici) sono scomparsi e sono stati sostituiti da zone stocastiche. Queste invadono lo spazio delle fasi quando si aumenta il parametro di stocasticità s facendo variare l'ampiezza dei campi magnetici o la velocità angolare del campo magnetico rotante. A causa della complessità dello spazio delle fasi a tre dimensioni, conviene considerarne delle sezioni, dette sezioni di Poincaré; una di esse costituisce la zona frontale dello spazio delle fasi qui rappresentato.

toria. Così, quando incliniamo l'ago rispetto al campo magnetico fisso e lo lasciamo a sé senza imprimergli una velocità iniziale, la traiettoria del punto rappresentativo del sistema è una curva chiusa simile a un'ellisse, tanto più grande quanto maggiore sarà l'angolo di scostamento dell'ago dalla direzione del campo magnetico; questo tipo di traiettoria corrisponde a un moto di oscillazione dell'ago intorno al campo magnetico fisso, e si dice allora che esso è imprigionato nel campo magnetico fisso. Se imprimiamo all'ago una spinta abbastanza forte da dargli un'energia sufficiente a superare l'angolo $\theta = \pi$ (la posizione in cui l'ago è orientato nella direzione opposta al campo magnetico), esso inizia a girare senza fine, accelerando quando la sua direzione si avvicina a quella del campo magnetico e rallentando quando se ne allontana. Nello spazio delle fasi, la traiettoria del suo punto rappresentativo è aperta (*curva in colore nell'illustrazione di pagina 37*). Una traiettoria molto particolare, che più avanti ricoprirà un ruolo fondamentale, è quella che costituisce il confine tra l'insieme delle traiettorie aperte e quello delle traiettorie chiuse: questa curva separatrice corrisponde alla situazione in cui la spinta impartita all'ago permette a quest'ultimo di raggiungere la sua posizione d'equilibrio instabile (in cui θ è uguale a π) con velocità nulla.

In questo problema integrabile, tutte le traiettorie sono regolari: l'effetto del campo magnetico fisso è rilevante per le traiettorie imprigionate, le quali si trovano in un dominio, delimitato dalle curve separatrici, che prende il nome di risonanza. Il cuore della risonanza è costituito da traiettorie a forma di ellisse, che corrispondono a piccole oscillazioni in cui l'ago è sottoposto a una forza di richiamo, creata dal campo magnetico, proporzionale al seno dell'angolo θ; dato che questo angolo è piccolo, la forza di richiamo in questa regione è in pratica proporzionale all'angolo θ stesso e il cuore della risonanza è dunque il dominio in cui le oscillazioni sono lineari, mentre il centro di questa regione è costituito esso stesso da un'ellisse infinitamente piccola: un punto ellittico. Analogamente, le cuspidi delle separatrici sono dette iperboliche perché in prossimità di questi punti le traiettorie sono iperboli che hanno per assi le separatrici. È per queste traiettorie che si fanno sentire in modo più rilevante gli effetti non lineari. In termini generali, l'incidenza di questi effetti cresce dal centro della risonanza verso le separatrici, come si può constatare in due modi. Per prima cosa le traiettorie a forma di ellisse, nel cuore della risonanza, si deformano nell'approssimarsi ai punti iperbolici, indicando in questo modo la comparsa di armoniche: le oscillazioni, allora, non sono più sinusoidali e compaiono anche frequenze multiple della frequenza fondamentale del moto. Inoltre, si perde l'isocronismo delle oscillazioni. (L'isocronismo, scoperto da Galileo per le oscillazioni lineari, consiste

Per costruire le sezioni di Poincaré si portano su uno stesso piano (*a destra*) i punti che si trovano all'intersezione delle diverse traiettorie con piani di «taglio» spaziati di 2π lungo l'asse ϕ. Per le traiettorie regolari che si avvolgono a elica sui tori di KAM, ciascuna intersezione è un punto posto su una curva (in questo caso una circonferenza): la sovrapposizione dei piani di taglio costituisce la curva punto per punto. Le traiettorie stocastiche, invece, corrispondono a moti disordinati dell'ago; i punti all'intersezione tra i piani di taglio e queste traiettorie si distribuiscono in una regione che si trova all'esterno degli ultimi tori di KAM.

nel fatto che la frequenza d'oscillazione non dipende dall'ampiezza dell'oscillazione stessa.) Per l'ago magnetico, la frequenza delle oscillazioni dipende dalla loro ampiezza; vicino al punto ellittico, dove le traiettorie sono ellissi quasi perfette e gli effetti non lineari sono di scarsa rilevanza, la frequenza d'oscillazione è massima e dipende poco dall'ampiezza. Via via che ci si avvicina alla separatrice, la frequenza diminuisce sempre più velocemente e diventa nulla in corrispondenza della separatrice. Aumentando l'intensità del campo magnetico fisso, l'estensione della risonanza lungo la direzione $\dot{\theta}$ cresce insieme alla frequenza d'oscillazione.

Tutto ciò che abbiamo appena osservato a proposito dell'ago rotante senza attrito in un campo magnetico fisso possiamo applicarlo al caso di un ago sottoposto solamente a un campo magnetico rotante, a condizione di immaginare che noi stessi ruotiamo insieme al campo magnetico: anche il campo rotante crea una risonanza, questa volta centrata sulla retta dello spazio delle fasi avente equazione $\dot{\theta} = \omega_0$. Ma che cosa avviene quando l'ago si trova collocato simultaneamente nei due campi magnetici?

La comparsa della stocasticità

Abbiamo detto in precedenza che, per questo sistema a due gradi di libertà un po' particolare, le traiettorie del punto rappresentativo del sistema si inscrivono nello spazio delle fasi a tre dimensioni individuato da $(\theta, \dot{\theta}, \phi)$. Ora, i problemi integrabili sono quelli le cui traiettorie s'inscrivono su tori di dimensione due: nel caso dell'ago posto nei due campi magnetici, se il problema fosse integrabile le traiettorie si inscriverebbero su superfici a due dimensioni nello spazio $(\theta, \dot{\theta}, \phi)$. Immaginiamo (e vedremo che, in certi casi, è quanto effettivamente avviene) che il moto dell'ago sia semplicemente la somma dei due moti corrispondenti al caso in cui vi sia uno solo dei due campi magnetici. Imprigionato nel campo fisso, l'ago oscillerebbe con frequenza F, e il punto rappresentativo del sistema descriverebbe un'ellisse nel piano $(\theta, \dot{\theta})$. Il campo magnetico rotante imprimerebbe all'ago un secondo moto d'oscillazione, con frequenza f_0 (uguale a $\omega_0/2\pi$), e il punto rappresentativo del sistema descriverebbe allora una traiettoria elicoidale giacente su un cilindro, dato che dovrebbe contemporaneamente ruotare nel piano $(\theta, \dot{\theta})$ e spostarsi a velocità costante lungo l'asse ϕ. (Si otterrebbe un toro rappresentando ϕ lungo una circonferenza.) Il passo dell'elica sarebbe allora uguale al rapporto F/f_0 delle frequenze di ciascun moto. In queste condizioni, le traiettorie stocastiche esistenti nel sistema si trovano in un volume che si può descrivere come una delle superfici del sistema supposto integrabile, che abbia assunto un certo «spessore»; in realtà, il sistema non integrabile possiede entrambi i tipi di soluzione: moti regolari corrispondenti ai cilindri nello spazio delle fasi e, come vedremo, anche moti stocastici; la natura del moto dipende dalle condizioni iniziali.

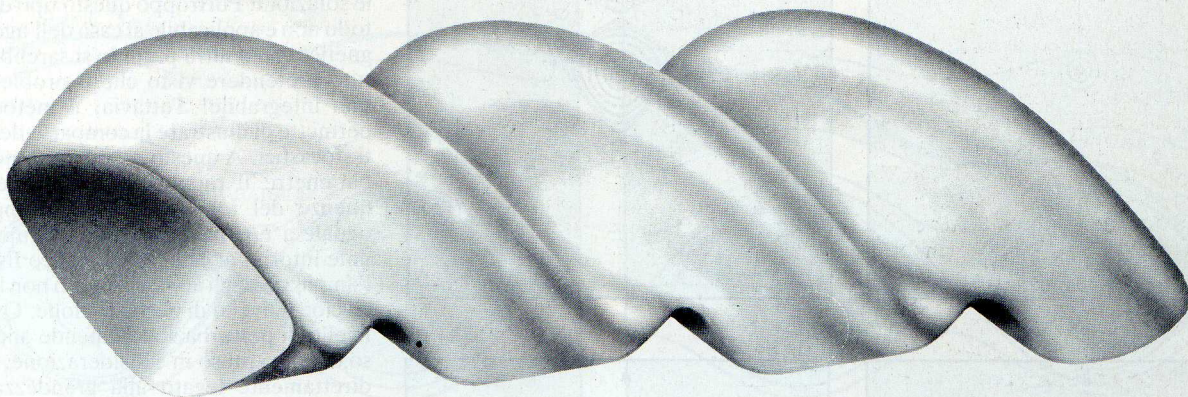
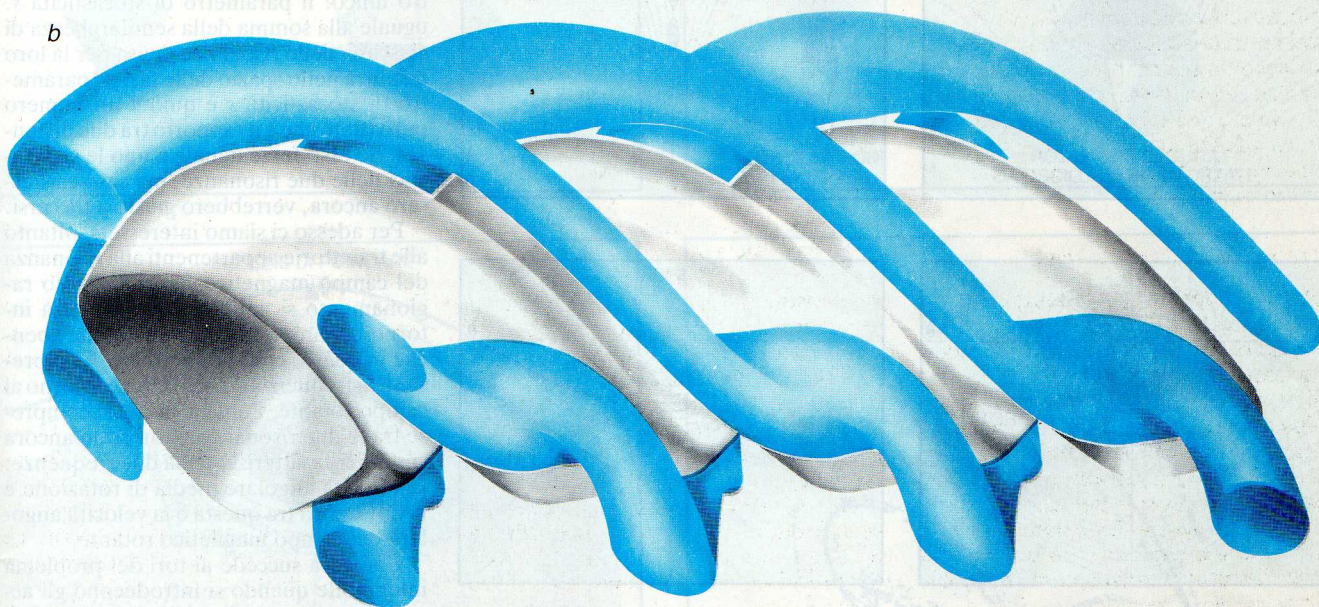
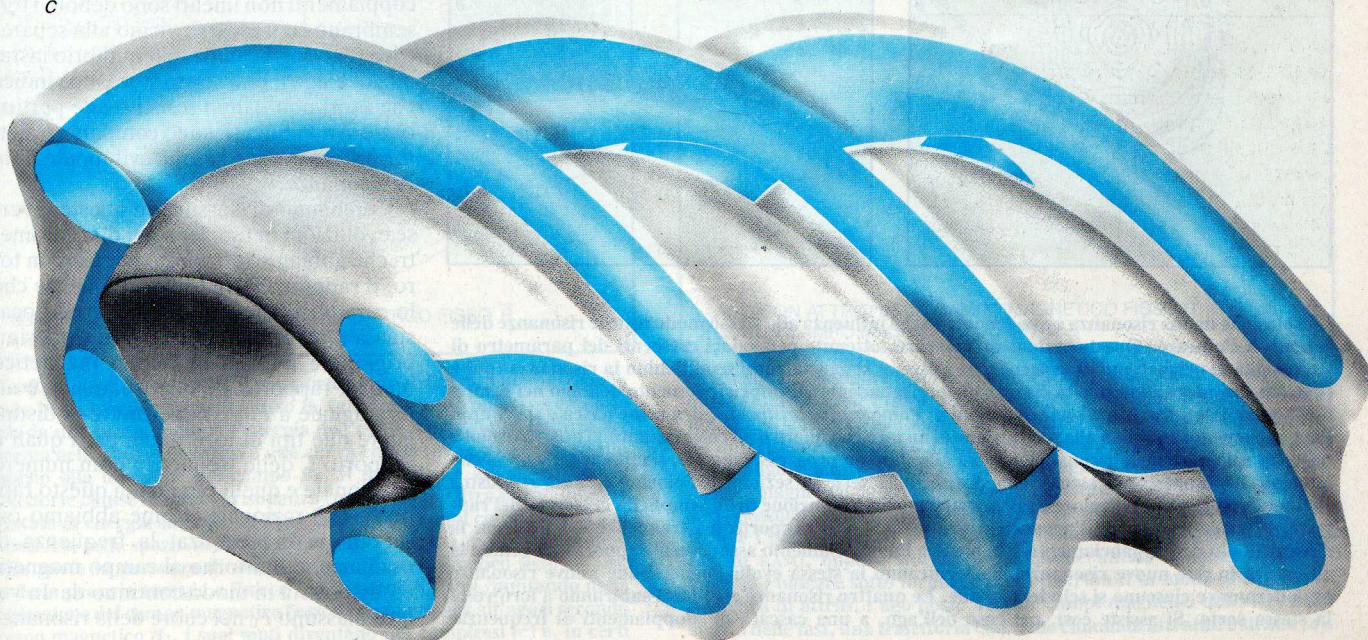
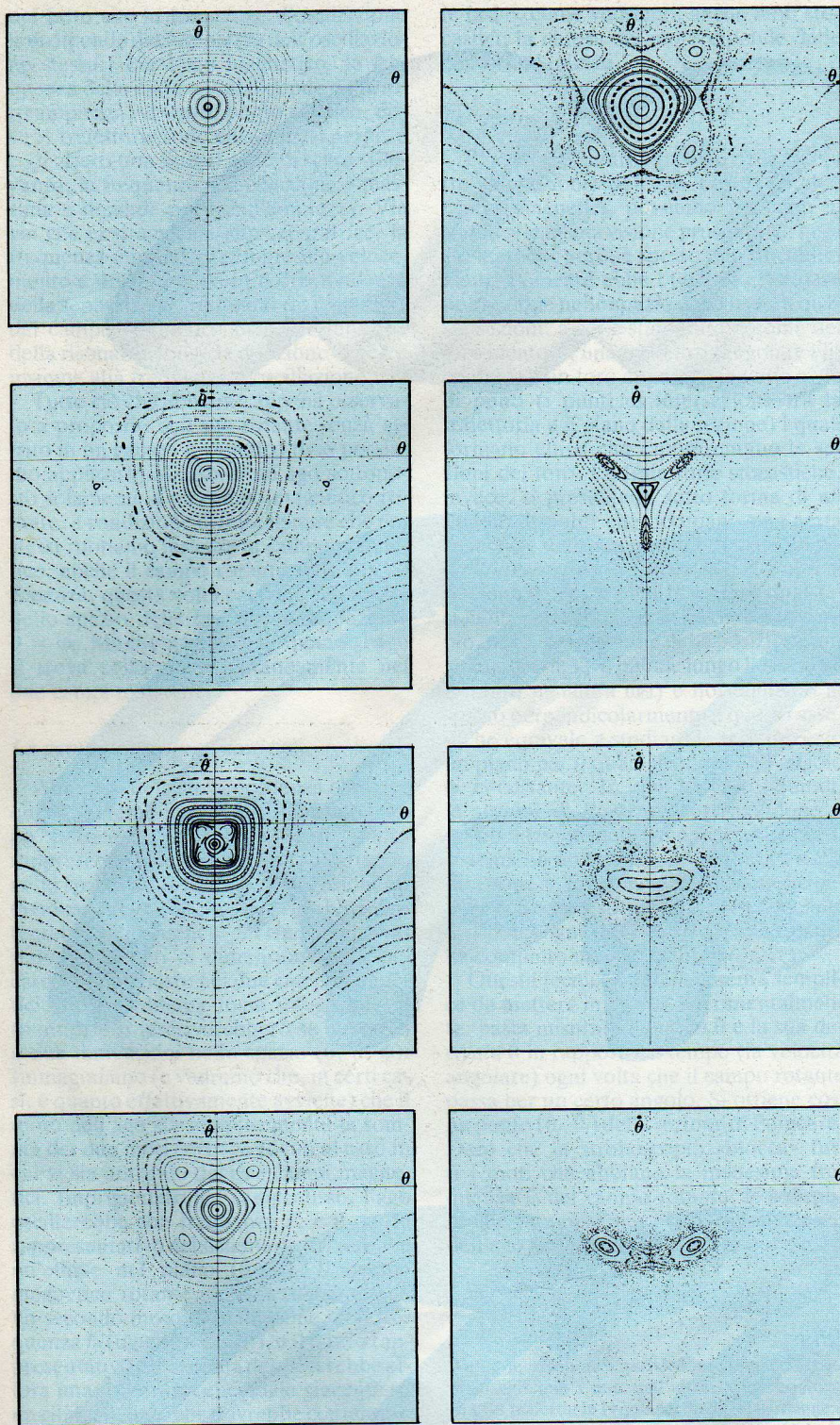
Le sezioni di Poincaré

Per rappresentare lo spazio delle fasi (a tre dimensioni) dell'ago posto nei due campi magnetici, possiamo tracciare le traiettorie in una visione prospettica; questo metodo, però, è spesso impraticabile. Henri Poincaré ebbe l'idea di effettuare una sezione nello spazio delle fasi. In queste sezioni, a cui è stato dato il nome del loro ideatore, una traiettoria regolare elicoidale su un toro appare come una serie di punti (i punti di intersezione tra la traiettoria e il piano della sezione) i quali formano ellissi che rappresentano le sezioni del toro. Le traiettorie stocastiche, invece, si presentano sotto forma di un nugolo di punti che occupano una certa superficie della sezione di Poincaré (*si veda l'illustrazione a pagina 38*). *A priori* si possono scegliere piani di sezione qualsiasi nello spazio delle fasi; nel nostro esempio, però, l'evoluzione della posizione angolare del taglio è banale lungo l'asse ϕ (ϕ è infatti uguale a $\omega_0 t$) e noi tagliamo lo spazio perpendicolarmente a questo asse, il che equivale a studiare le traiettorie in un piano per il quale il valore dell'angolo ϕ è costante. In questo caso, quindi, le sezioni di Poincaré corrispondono a un'operazione molto semplice: dato che il sistema ha un'evoluzione periodica nella direzione ϕ, fare una sezione di questo genere equivale a sottoporre a stroboscopia il sistema ogni volta che il campo rotante ha compiuto un giro completo.

Questa tecnica è relativamente semplice da mettere in pratica sperimentalmente: basta misurare l'angolo θ e la sua derivata $\dot{\theta}$ in rapporto al tempo (la velocità angolare) ogni volta che il campo rotante passa per un certo angolo. Si ottiene così un punto $(\theta_i, \dot{\theta}_i)$ della sezione di Poincaré. Dato che la stroboscopia «blocca» tutti i moti che abbiano la medesima frequenza f_0 del campo rotante, le soluzioni che rappresentano l'imprigionamento dell'ago nel campo rotante (in questo caso, l'ago segue globalmente il campo rotante) appariranno, nello spazio delle fasi, come le soluzioni in cui l'ago è imprigionato nel campo fisso, ma con uno scostamento di ω_0 lungo l'asse $\dot{\theta}$. Le sezioni di Poincaré comportano dunque due risonanze principali, una associata al campo magnetico fisso e l'altra associata al campo magnetico rotante. Queste risonanze, però, sembrano incomplete: solo le traiettorie centrate intorno ai punti ellittici appaiono chiaramente (in effetti sono rimaste regolari), mentre le traiettorie vicine alle separatrici si sono mescolate dando luogo a uno strato stocastico. Inoltre sono apparse in modo strutturato nuove risonanze che si dispongono intorno alle prime e le riproducono a scala ridotta.

Per capire a che cosa corrispondano queste strutture, dobbiamo tornare al caso integrabile: se il sistema fosse integrabile le traiettorie, nello spazio delle fasi, si inscriverebbero su tori a due dimensioni generati da due moti d'oscillazione. Si potrebbe descrivere ciascuno di questi tori con le due frequenze fondamentali delle sue oscillazioni; così, quando l'ago è imprigionato nel campo magnetico fisso, queste due frequenze sono la frequenza del campo magnetico rotante f_0 e la frequenza d'oscillazione F dell'ago intorno al campo magnetico fisso. Se il sistema fosse integrabile, questi due oscillatori sarebbero indipendenti e i loro moti sarebbero completamente separati. Nel caso reale dell'ago, gli oscillatori non sono più indipendenti: gli accoppiamenti non lineari distruggono alcuni dei tori che abbiamo appena esaminato. Per descrivere lo spazio delle fasi in questo caso non integrabile conviene considerare i tori del caso immaginario (integrabile) e seguirne l'evoluzione «innestando» progressivamente gli accoppiamenti non lineari.

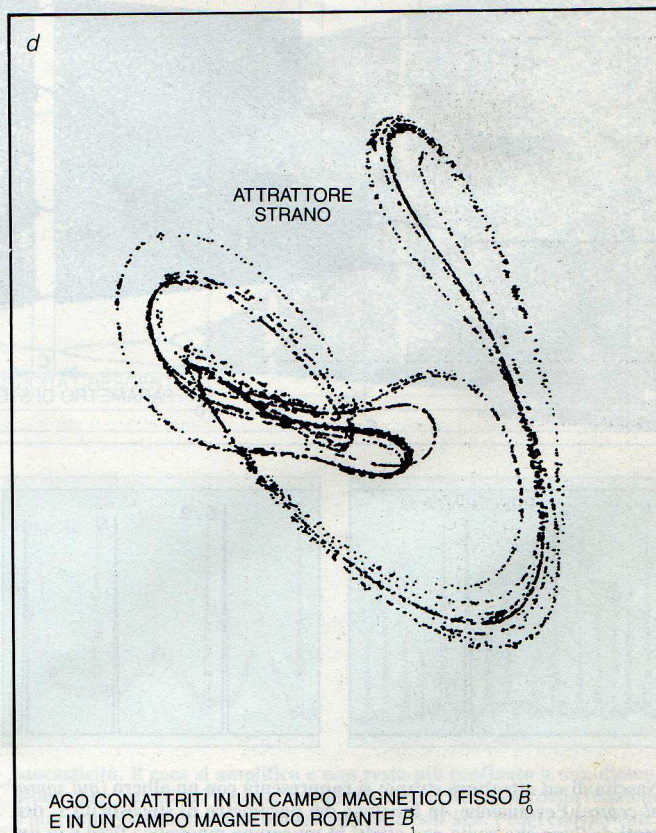
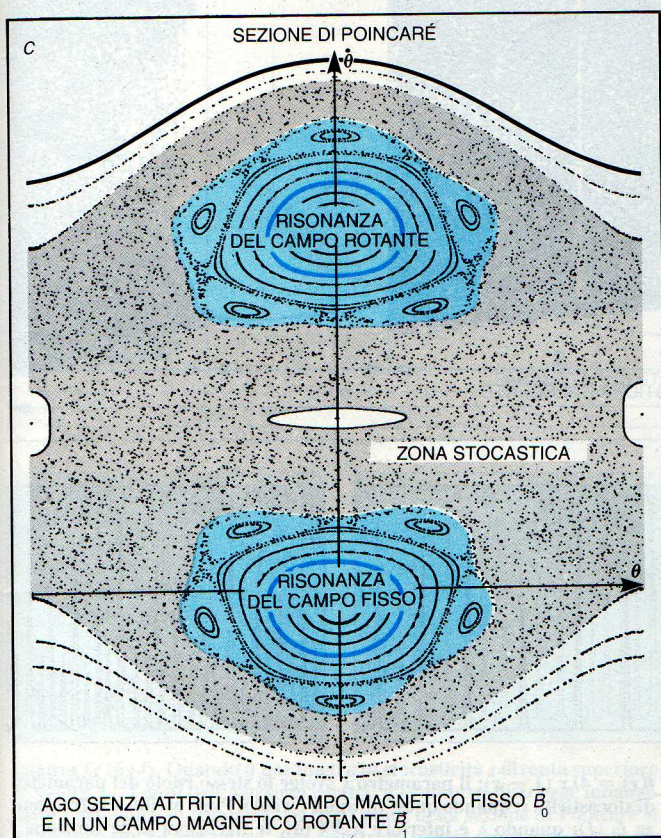
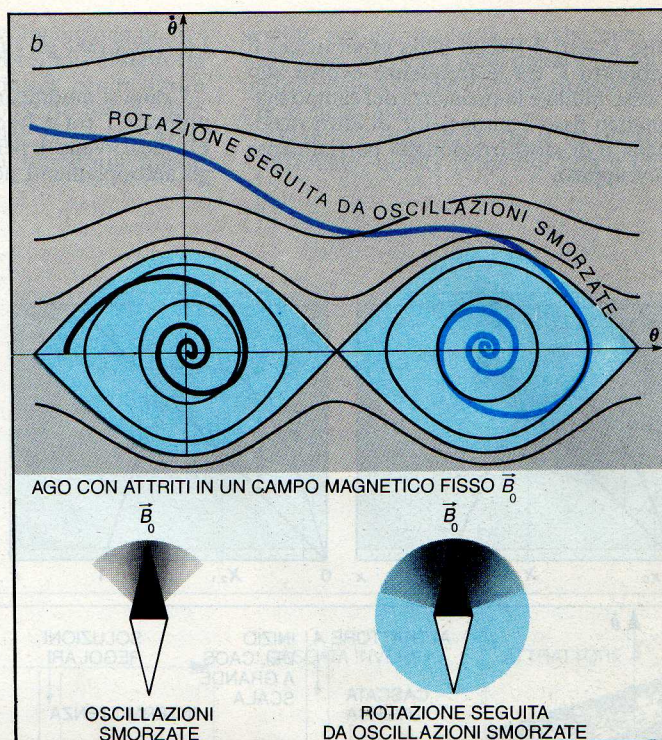
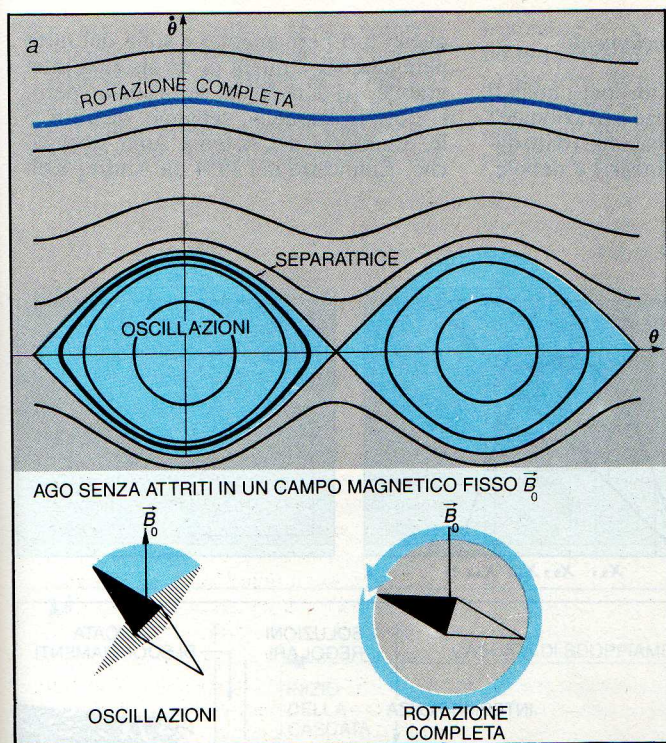
In certo qual modo, utilizziamo implicitamente un metodo di perturbazione: per ottenere le soluzioni di un problema incognito, si considerano quelle di un sistema risolvibile molto simile, e la differenza tra i due sistemi costituisce la perturbazione che si fa crescere in modo continuo per passare da un problema all'al-

Tutte le risonanze possiedono una struttura paragonabile a quella delle bambole russe che si inseriscono l'una nell'altra. Ciascun toro di KAM (*a*) è generato dall'avvolgimento a elica di una traiettoria regolare. Nell'illustrazione è rappresentato solo il toro su cui si avvolge la traiettoria, e non la traiettoria stessa. Queste traiettorie regolari corrispondono a moti dell'ago associati a due frequenze e alle loro armoniche e in cui il rapporto tra le due frequenze fondamentali sia un numero irrazionale. I tori che sarebbero generati dall'avvolgimento di traiettorie associate a due frequenze commensurabili (il cui rapporto sarebbe allora un numero razionale p/q) non esistono perché il problema non è integrabile; questi tori sono sostituiti da zone stocastiche che circondano degli «agganciamenti», cioè piccole risonanze (*in colore*) centrate (*b*) intorno a q punti ellittici (qui q = 4). Nelle sezioni di Poincaré questi agganciamenti appaiono come piccole risonanze distribuite intorno alle risonanze principali. Il modo di costruzione delle sezioni di Poincaré spiega perché le sezioni di questi piccoli tori di KAM corrispondano in realtà a un'unica traiettoria regolare. Il moto associato a questa traiettoria regolare comporta due frequenze fondamentali e le loro armoniche: una di queste frequenze è la frequenza d'oscillazione dell'ago ($f_0/4$) e l'altra è la frequenza dell'oscillazione di fase che caratterizza l'agganciamento. Dato che ciascun numero reale razionale è compreso tra numeri reali irrazionali, gli agganciamenti e la zona stocastica che hanno sostituito un toro razionale sono compresi in un altro toro di KAM (il toro esterno mostrato in *c*) associato a un moto dell'ago che comporta due frequenze fondamentali il cui rapporto è un numero irrazionale. Questo incastro di tori si ripete qualunque sia la scala del disegno.

a
b
c

L'esplosione di una risonanza è uno fenomeno che influenza allo stesso modo tutte le risonanze delle sezioni di Poincaré. Queste sezioni di Poincaré, ottenute per valori crescenti del parametro di stocasticità *s*, permettono di vedere come i diversi agganciamenti modifichino la risonanza (qui è presa in considerazione solo la risonanza corrispondente all'imprigionamento dell'ago nel campo magnetico fisso). La prima serie di sezioni di Poincaré (*a sinistra*) mostra come nasca e si sviluppi l'agganciamento associato alla sostituzione di un toro con quattro piccole risonanze: i quattro isolotti nascono nel cuore della risonanza e se ne allontanano progressivamente, ampliandosi via via che aumenta il parametro di stocasticità. La seconda serie di sezioni di Poincaré (*a destra*) illustra l'importanza sempre maggiore degli agganciamenti in funzione della semplicità crescente del rapporto Γ tra le frequenze: l'agganciamento associato al rapporto tra frequenze 1/3 deforma la risonanza più dell'agganciamento 1/4, mentre l'agganciamento associato al rapporto 1/2 scinde la risonanza in due nuove risonanze che seguiranno la stessa evoluzione. Le due nuove risonanze esploderanno e ciascuna si scinderà in due. Le quattro risonanze risultanti subiranno a loro volta la stessa sorte. Si assiste così, nel caso dell'ago, a una cascata di sdoppiamenti di frequenza.

tro, supponendo che lo stesso avvenga per le soluzioni. Purtroppo questo tipo di metodo non è applicabile al caso dell'ago magnetico, né d'altra parte lo si sarebbe potuto pretendere visto che il problema è non integrabile! Tuttavia, il metodo ci permette di illustrare la comparsa dei moti stocastici. A questo scopo useremo due parametri: il rapporto Γ delle due frequenze del toro considerato (rapporto uguale a F/f_0 per le traiettorie imprigionate intorno al campo magnetico fisso) e l'ampiezza degli accoppiamenti non lineari, cioè il livello di perturbazione. Questo livello di perturbazione dipende anch'esso dal toro preso in considerazione, ma è direttamente legato alla grandezza dei due campi magnetici e alla frequenza del campo magnetico rotante. Questi parametri vengono raggruppati in un parametro unico: il parametro di stocasticità *s*, uguale alla somma della semilarghezza di ciascuna delle risonanze divisa per la loro distanza nello spazio delle fasi. Il parametro di stocasticità *s* è quindi un numero puro (dato che è il rapporto tra due distanze). Se esso fosse uguale a uno le separatrici delle due risonanze, nel caso esistessero ancora, verrebbero giusto a toccarsi.

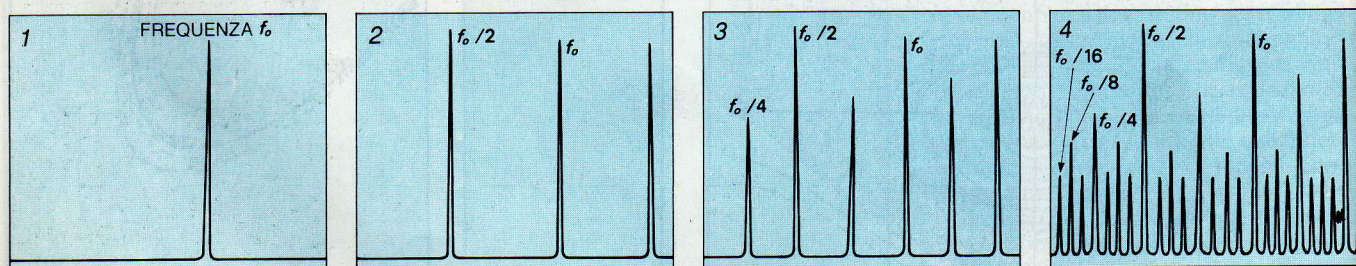
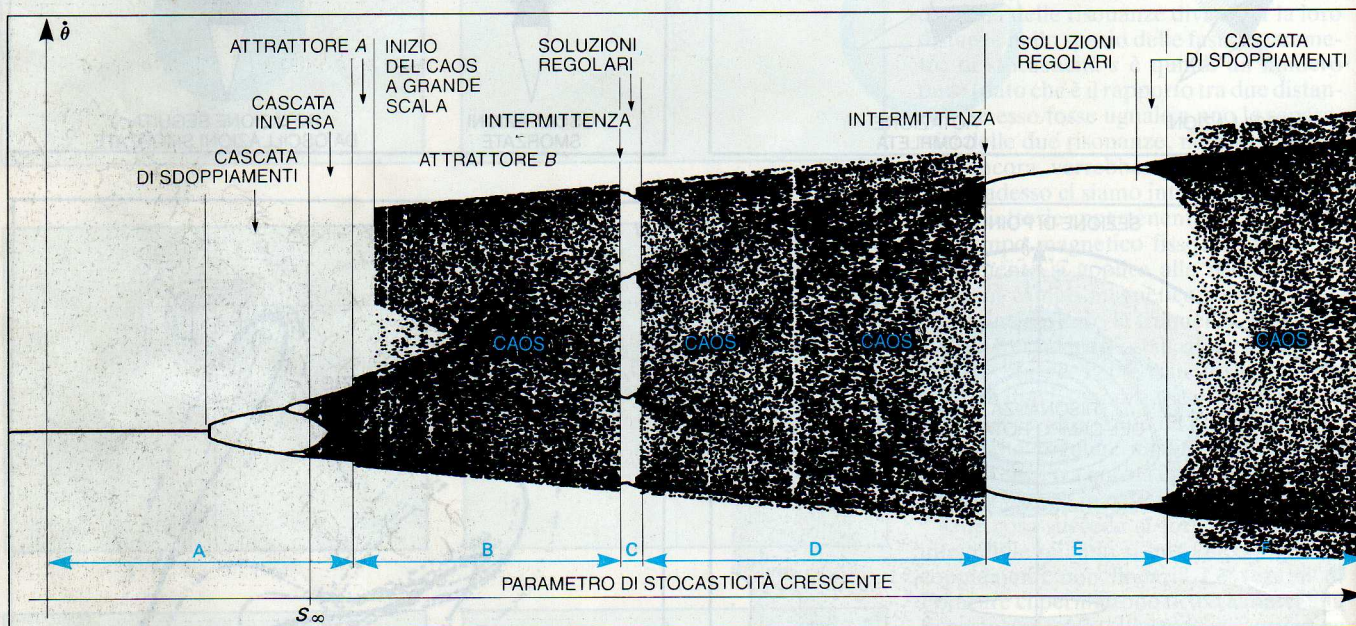
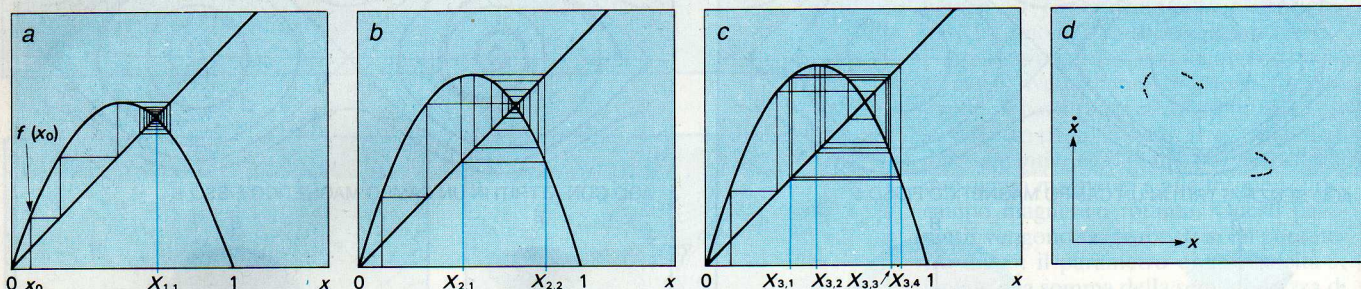
Per adesso ci siamo interessati soltanto alle traiettorie appartenenti alla risonanza del campo magnetico fisso; lo stesso ragionamento si applica alla risonanza intorno al campo magnetico rotante, benché in questo caso la frequenza *F* rappresenti la frequenza d'oscillazione intorno al campo rotante. Per le traiettorie comprese tra le due risonanze, i tori sono ancora superfici caratterizzate da due frequenze: la velocità angolare media di rotazione e la differenza tra questa e la velocità angolare del campo magnetico rotante.

Che cosa succede ai tori del problema integrabile quando si introducono gli accoppiamenti non lineari? Le sezioni di Poincaré ci permettono di avvicinarci alla risposta (*si veda l'illustrazione a pagina 38*): vicino al punto ellittico, dove gli accoppiamenti non lineari sono deboli, i tori sembrano conservarsi; vicino alla separatrice (che corrisponde all'equilibrio instabile dell'ago), uno strato stocastico indica che i tori non hanno resistito alle perturbazioni non lineari; sono comparse traiettorie stocastiche e si sono sviluppate nuove risonanze.

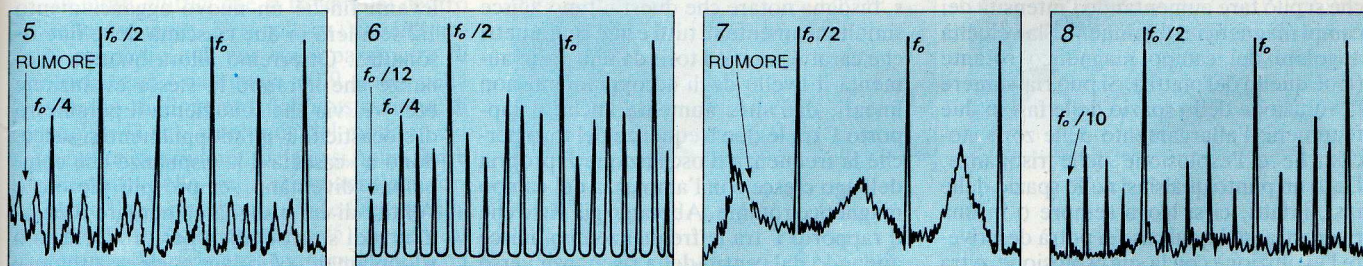
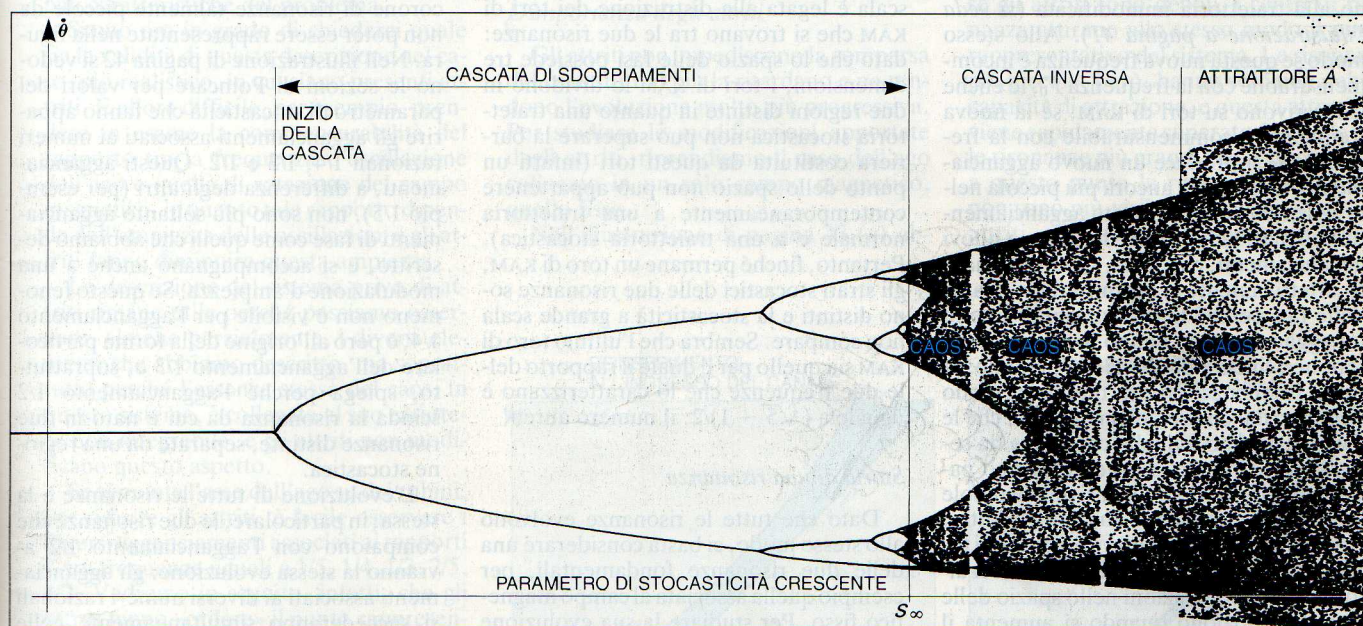
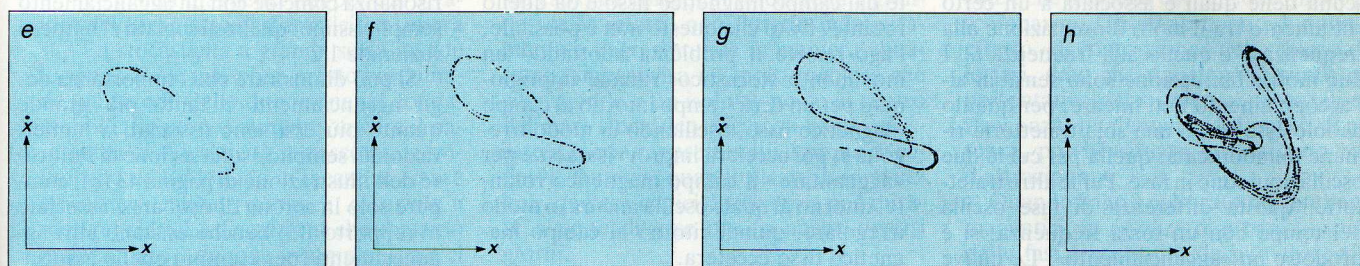
Cerchiamo di classificare queste diverse evoluzioni in funzione dei due parametri che abbiamo introdotto per ciascun toro: il rapporto Γ tra le due frequenze che lo caratterizzano e il parametro di stocasticità *s*. Consideriamo le traiettorie imprigionate intorno al campo magnetico fisso: il rapporto Γ delle frequenze è allora uguale a F/f_0 e ci permette di distinguere due tipi di tori: quelli per i quali il rapporto Γ delle frequenze è un numero irrazionale e quelli per i quali questo rapporto Γ è razionale. Come abbiamo osservato in precedenza, la frequenza di oscillazione *F* intorno al campo magnetico fisso varia in modo continuo da un valore massimo F_0 nel cuore della risonanza

Gli attriti modificano molto il comportamento dell'ago. Quando questo ruota in un campo magnetico fisso B_0 (*a*), il suo moto può essere di due tipi: da una parte, quando l'energia dell'ago è scarsa, la traiettoria del sistema nello spazio delle fasi è una curva chiusa (*in nero*) e l'ago oscilla intorno alla direzione del campo B_0. Il dominio in cui si trovano le traiettorie chiuse si chiama risonanza (*in colore*). D'altra parte, quando è dotato di energia sufficiente, l'ago effettua rotazioni complete e la sua traiettoria nello spazio delle fasi è una curva aperta (*in colore*). Quando l'ago ruota in presenza di attriti in un campo magnetico fisso B_0 (*b*), perde la propria energia e finisce con l'immobilizzarsi, puntando nella in direzione del campo magnetico fisso B_0. Se si applica all'ago il secondo campo magnetico B_1, i suoi moti diventano più complessi (*c*) e, in certi casi, caotici. Quando l'ago ruota senza attriti, lo spazio delle fasi è a tre dimensioni (*si veda l'illustrazione a pagina 38*) e per descrivere i moti dell'ago si ricorre alle sezioni di Poincaré. Queste sezioni comprendono due risonanze principali (*in colore*) che corrispondono ai moti per i quali l'ago è imprigionato nel campo magnetico fisso (*in basso*) o nel campo magnetico rotante (*in alto*). Nelle sezioni di Poincaré le traiettorie possono essere regolari, e in questo caso le intersezioni tra le traiettorie e i piani di taglio si distribuiscono su curve chiuse (le sezioni dei tori di KAM), oppure stocastiche, e queste ultime appaiono nelle sezioni di Poincaré sotto forma di punti distribuiti in modo aleatorio. In presenza di attriti, l'ago posto nei due campi magnetici descrive, nello spazio delle fasi, una traiettoria che viene chiamata attrattore strano (*d*).

fino a zero a livello della separatrice; il rapporto Γ tra le frequenze evolve allo stesso modo e la risonanza del campo magnetico fisso è un insieme di «tori razionali» e di «tori irrazionali» parzialmente sovrapposti.

I tori di KAM e gli «agganciamenti»

Come si modificano i tori per i quali il rapporto Γ tra le frequenze è irrazionale? Finché il livello di perturbazione (dato dagli accoppiamenti non lineari) è debole, questi tori permangono e sono del tutto paragonabili a quelli di un sistema integrabile. Al di là di una certa soglia, però, i «tori irrazionali» vengono distrutti e le traiettorie del sistema sono stocastiche. Enunciato nel 1954 da Andrej Kol-

Nascita di un attrattore strano: si rappresenta con un albero (*qui sopra al centro*) l'evoluzione, in funzione del parametro di stocasticità s, dei moti dell'ago che ruota con attriti in un campo magnetico fisso e in un campo magnetico rotante a velocità angolare costante ω_0. Per ciascuno dei valori del parametro di stocasticità s dati in ascissa, si riportano in ordinata le differenti velocità angolari dell'ago sottoposto a stroboscopia alla frequenza f_0 uguale a $\omega_0/2\pi$. Quando si trova un solo punto per un valore dato del parametro di stocasticità s, l'ago oscilla solo alla frequenza f_0 e alle sue armoniche; quando si trovano due punti, l'ago oscilla alle frequenze f_0 e $f_0/2$ e alle loro armoniche. Un insieme continuo di punti rappresenta un moto caotico dell'ago. La regione A corrisponde a una cascata di sdoppiamenti di periodo seguita da una cascata inversa (*si veda nella pagina a fronte un ingrandimento di questa regione dell'albero*). Durante questa fase, si può realizzare un modello delle evoluzioni del sistema servendosi delle iterazioni di una funzione a una variablie come $f(x) = 4\lambda x(1-x)$; il parametro λ svolge lo stesso ruolo del parametro di stocasticità s: quando quest'ultimo è piccolo, l'ago oscilla alla frequenza f_0 (*1*); quando λ è inferiore a 3/4 (*a*), le iterazioni della funzione f giustificano il fatto che il moto abbia una sola frequenza: le iterazioni della funzione conducono a un unico valore $X_{1,1}$. Quando il parametro di stocasticità s aumenta, avviene il primo sdoppiamento, che si traduce nella comparsa, nello spettro di Fourier (*2*), dell'armonica di frequenza $f_0/2$. Lo sdoppiamento appare anche per le iterazioni della funzione f quando λ è superiore a 3/4 e inferiore a 0,86... (*b*): i valori successivi della funzione f si distribuiscono intorno ai due valori $X_{2,1}$ e $X_{2,2}$. Se si aumenta ancora il parametro di stocasticità s, si produce un nuovo sdoppiamento, la frequenza $f_0/4$ appare nel moto dell'ago (*3*) e i valori consecutivi della funzione f si distribuiscono intorno a quattro valori $X_{3,1}$, $X_{3,2}$, $X_{3,3}$ e $X_{3,4}$ (*c*). Il fenomeno si ripete all'infinito, ma si possono osservare sperimentalmente gli sdoppiamenti solo fino alla comparsa dell'armonica di fre-

mogorov, questo risultato apparentemente semplice venne dimostrato da Vladimir I. Arnold e Jürgen Moser nel 1962: costituisce il teorema KAM e i tori non distrutti hanno preso il nome di tori di KAM. Per questi tori, i moti d'oscillazione alle frequenze F e f_0 si sovrappongono in modo indipendente; inoltre, una traiettoria su un toro di KAM finisce per ricoprire quest'ultimo in modo denso (ossia, data una qualsiasi superficie di questo toro, piccola a piacere, la traiettoria finisce sempre per attraversarla) perché questa traiettoria è un'elica il cui passo è incommensurabile con 2π e che non ripassa mai per lo stesso punto.

Quando il rapporto Γ tra le frequenze è razionale, questa proprietà non è più

quenza $f_0/16$ (4). Quando il parametro di stocasticità s diventa superiore al valore s_∞, i moti dell'ago sono caotici: il «rumore» distrugge innanzitutto l'armonica di frequenza $f_0/16$ e i punti della sezione di Poincaré si distribuiscono su una curva. Per realizzare questa sezione di Poincaré, è stata registrata la forza elettromotrice indotta dai moti dell'ago in bobine di rilevamento; questo segnale e la sua derivata alla frequenza f_0 sono stati sottoposti a stroboscopia e si sono riportate in grafico le differenti coppie (x_i, $\dot{x}_i$). Aumentando ancora il parametro di stocasticità s, si osserva una cascata inversa: le armoniche scompaiono nell'ordine inverso a quello della loro comparsa. In 5 è mostrata la distruzione dell'armonica di frequenza $f_0/8$ e, in occasione di questa distruzione, si osservano gli agganciamenti 1/5 e 1/3 corrispondenti alle frequenze $f_0/20$ e $f_0/12$. Il fenomeno si ripete per $f_0/4$ e si vede, in 6, come il caos tenda a distruggere questa armonica con la comparsa dell'agganciamento 1/5 alla frequenza $f_0/10$ (8). A partire da questo valore del parametro di stocasticità, il caos si amplifica e non resta più confinato a una dimensione nelle sezioni di Poincaré; a questo punto le iterazioni della funzione f non rappresentano più un modello dei moti dell'ago. Progressivamente, invece, si costituisce un attrattore strano (*e*, *f*, *g*, *h*). L'attrattore *g* corrisponde a oscillazioni irregolari dell'ago imprigionato nel campo magnetico fisso. Nella regione *B* dell'albero, il caos è a grande scala e l'ago non è più imprigionato (*h*). Dopo la regione *B* compare una fase regolare (*regione C*), e il passaggio dalla regione *B* alla *C* avviene attraverso una transizione del tipo dell'intermittenza (*si veda l'illustrazione a pagina 47*), mentre il passaggio dalla *C* alla *D*, o quello dalla *E* alla *F*, avviene attraverso una cascata di sdoppiamenti di frequenza. La regione *E* corrisponde a una nuova fase regolare e la restante parte dell'albero è un'alternanza di fasi regolari e di fasi stocastiche. Questi risultati sono stati ottenuti al CEA di Gif-sur-Yvette; C. Poitou, M. Labouise, B. Ozenda e M. Clément hanno progettato e realizzato gli elementi meccanici.

verificata: infatti, se il rapporto Γ tra le frequenze è per esempio uguale a p/q (dove p e q sono numeri interi), dopo q giri completi del campo rotante l'elica si richiude su se stessa e il moto si ripete esattamente identico. La traiettoria è allora una curva chiusa sul toro e non ricopre più il toro stesso. In questo caso il toro è costituito da un'infinità di queste curve, ciascuna delle quali è associata a un certo sfasamento tra il moto d'oscillazione alla frequenza F e quello alla frequenza f_0. I due moti d'oscillazione sono sensibili all'accoppiamento non lineare, per quanto debole esso sia, e una sola traiettoria rimane immodificata: quella per cui le due oscillazioni sono in fase. Per le altre traiettorie, questa differenza di fase oscilla nel tempo con un terza frequenza: si è prodotto un «agganciamento». Le nuove traiettorie sono eliche che ruotano intorno alla traiettoria immodificata (*si veda l'illustrazione a pagina 41*). Allo stesso modo se questa nuova frequenza è incommensurabile con la frequenza F_0, le eliche si inscrivono su tori di KAM; se la nuova frequenza è commensurabile con la frequenza f_0 si produce un nuovo agganciamento, a una scala ancora più piccola nello spazio delle fasi. Questo agganciamento può dare luogo a tori di KAM o a nuovi agganciamenti, e così via. Questo è quanto effettivamente si produce nello spazio delle fasi perché le oscillazioni di fase in un agganciamento hanno lo stesso carattere non lineare di quelle del pendolo; la terza frequenza varierà quindi in modo continuo dal centro della risonanza che le è associata fino a quella che funge da separatrice. Complessivamente, si può paragonare lo spazio delle fasi alle bambole russe che si inseriscono una dentro l'altra.

Questo ragionamento vale per tutti i tori dello spazio delle fasi. Occupiamoci allora delle modificazioni nello spazio delle fasi che appaiono quando si aumenta il valore del parametro di stocasticità s, cosa che si può fare aumentando l'intensità dei campi magnetici o diminuendo la velocità angolare del campo magnetico rotante (cioè quello del piatto). Si può riassumere l'evoluzione dello spazio delle fasi in due fenomeni: l'allargamento delle zone stocastiche e l'esplosione delle risonanze. Dato un punto qualsiasi nello spazio delle fasi, infatti, ci si trova sempre o in una risonanza, nel qual caso si potrà descrivere l'evoluzione con la sua esplosione, o tra due risonanze, e l'evoluzione sarà allora determinata dall'allargamento delle zone stocastiche che le contornano.

La stocasticità a grande scala

Le risonanze associate ai campi magnetici fisso e rotante sono contornate da zone stocastiche che hanno sostituito le separatrici; se si aumenta il parametro di stocasticità s, queste zone stocastiche si ispessiscono e finiscono per fondersi dando luogo alla stocasticità a grande scala. È molto semplice descrivere questo fenomeno prendendo in considerazione il criterio di sovrapposizione delle risonanze: per un parametro di stocasticità uguale a uno le separatrici delle risonanze, se esistessero ancora, verrebbero a toccarsi; per valori del parametro di stocasticità superiori a uno i dominî di imprigionamento delimitati dalle separatrici si sovrapporrebbero e l'ago della bussola sarebbe imprigionato contemporaneamente dal campo magnetico fisso e da quello rotante. Dato che questo non è possibile, l'ago risolve il problema adottando un movimento stocastico: rimane imprigionato per un certo tempo intorno al campo magnetico fisso, oscillando in modo irregolare, poi accelera improvvisamente per «agguantare» il campo magnetico rotante, intorno al quale oscilla ancora in modo irregolare, quindi ritorna al campo magnetico fisso eccetera.

La comparsa delle stocasticità a grande scala è legata alla distruzione dei tori di KAM che si trovano tra le due risonanze: dato che lo spazio delle fasi possiede tre dimensioni, i tori di KAM lo dividono in due regioni distinte in quanto una traiettoria stocastica non può superare la barriera costituita da questi tori (infatti un punto dello spazio non può appartenere contemporaneamente a una traiettoria normale e a una traiettoria stocastica). Pertanto, finché permane un toro di KAM, gli strati stocastici delle due risonanze sono distinti e la stocasticità a grande scala non compare. Sembra che l'ultimo toro di KAM sia quello per il quale il rapporto delle due frequenze che lo caratterizzano è uguale a $(\sqrt{5} - 1)/2$: il numero aureo!

Storia di una risonanza

Dato che tutte le risonanze evolvono allo stesso modo, ci basta considerare una delle due risonanze fondamentali, per esempio quella associata al campo magnetico fisso. Per studiare la sua evoluzione in funzione del parametro di stocasticità s, bisogna notare che quest'ultimo agisce simultaneamente su tutti e due i parametri che caratterizzano i tori: da una parte aumenta il livello degli accoppiamenti non lineari, dall'altra aumenta anche il rapporto Γ tra le due frequenze del toro perché la frequenza d'oscillazione F propria dell'ago cresce con l'ampiezza del campo magnetico rotante. Abbiamo già visto che il rapporto Γ tra le frequenze diminuisce andando dal centro della risonanza verso la regione stocastica che sostituisce le separatrici, e che la posizione di un agganciamento p/q è determinata dal punto in cui il rapporto Γ è uguale a p/q. Quando si aumenta il parametro di stocasticità s, la «corona» di q isolotti corrispondenti all'agganciamento p/q si allontana dal centro della risonanza verso la regione stocastica. Un agganciamento nasce al centro della risonanza perché questo è il primo punto in cui il rapporto Γ tra le frequenze assume il valore p/q quando aumenta il parametro di stocasticità s. Si assiste così a una successione di agganciamenti corrispondenti a rapporti Γ tra frequenze sempre più grandi, dal cuore della risonanza fino alla zona stocastica che la circonda.

Preannunciando l'esplosione delle risonanze abbiamo anticipato la loro storia: proprio come la nascita della stocasticità a grande scala è legata alla distruzione dell'ultimo toro di KAM per il quale il rapporto Γ tra le frequenze è il numero più irrazionale possibile, così l'esplosione di una risonanza coincide con un agganciamento semplicissimo: quello associato al numero razionale 1/2.

Si può dimostrare che l'importanza degli agganciamenti è tanto più grande quanto più essi sono associati a numeri razionali semplici: sulla sezione di Poincaré dell'illustrazione di pagina 43 (*c*), compare solo la corona di risonanze associate al rapporto 1/5, benché esistano altri agganciamenti (per esempio quello associato al rapporto 6/35) che danno luogo a corone di risonanze talmente piccole da non poter essere rappresentate nella figura. Nell'illustrazione di pagina 42 si vedono le sezioni di Poincaré per valori del parametro di stocasticità che fanno apparire gli agganciamenti associati ai numeri razionali 1/4, 1/3 e 1/2. Questi agganciamenti, a differenza degli altri (per esempio 1/5), non sono più soltanto agganciamenti di fase come quelli che abbiamo descritto, e si accompagnano anche a una modulazione d'ampiezza. Se questo fenomeno non è visibile per l'agganciamento 1/4, è però all'origine della forma particolare dell'agganciamento 1/3 e, soprattutto, spiega perché l'agganciamento 1/2 scinda la risonanza da cui è nato in due risonanze distinte, separate da una regione stocastica.

L'evoluzione di tutte le risonanze è la stessa; in particolare, le due risonanze che compaiono con l'agganciamento 1/2 avranno la stessa evoluzione: gli agganciamenti associati ai diversi numeri razionali si succederanno simultaneamente nelle due risonanze, diventando sempre più rilevanti finché un nuovo agganciamento 1/2 scinderà in due ciascuna delle due risonanze. Otterremo allora quattro risonanze che avranno la stessa evoluzione ecc. Via via che si aumenta il parametro di stocasticità s, gli sdoppiamenti si succedono a «cascata»; le risonanze che compaiono diventano sempre più piccole e l'entità di cui bisogna aumentare il parametro di stocasticità s per produrre uno sdoppiamento si riduce anch'essa sempre di più: se si indica con s_n il valore del parametro per il quale compare l'n-esimo sdoppiamento, la successione dei valori s_n tende verso un limite s_∞ in cui si è prodotta un'infinità di sdoppiamenti.

Questa cascata di sdoppiamenti, per quanto di natura differente da quella che si osserva quando si itera la funzione $f(x) = 4\lambda x(1 - x)$, possiede le stesse proprietà di universalità introdotte da Mitchell Feigenbaum. Così, per esempio, i rapporti tra le distanze successive delle soglie di sdoppiamento $(s_n - s_{n-1})/(s_{n+1} - s_n)$ tendono verso un valore limite δ, identico per tutti i problemi di una stessa classe:

tutti i sistemi meccanici a due gradi di libertà e senza attriti presentano delle cascate di sdoppiamenti con la stessa costante δ, indipendentemente dalla natura di questi sistemi. Per questa classe di problemi, la costante δ è uguale a 8,72...

I moti di un ago reale

Che cosa succede quando si impongono all'ago campi magnetici tali per cui il valore del parametro di stocasticità s supera s_∞? La stocasticità a grande scala invade lo spazio delle fasi e ben presto non rimane più traccia delle due risonanze fondamentali. Se però si aumenta ancora il parametro di stocasticità s, nascono, si sviluppano ed esplodono in mezzo al «mare stocastico» nuove risonanze, associate a moti regolari. La descrizione che abbiamo appena fatto è dunque incompleta, ma ci ha permesso di mettere in luce le principali caratteristiche del problema.

Siamo ora in grado di chiederci quale sia la validità di questa descrizione nel caso, più realistico, in cui siano presenti attriti. È allora difficile, per esempio, prendere in esame la commensurabilità del rapporto tra la frequenza d'oscillazione dell'ago e quella di rotazione del campo magnetico, in quanto tale rapporto dipende dall'ampiezza delle oscillazioni e gli attriti fanno diminuire questa ampiezza.

La descrizione del sistema privo di attriti rimane utile perché possiamo osservare, almeno parzialmente, i diversi elementi che abbiamo descritto, ma soprattutto perché l'essenza stessa del caos, in questo sistema, è collegata al suo carattere non integrabile, e gli attriti non modificano questo aspetto.

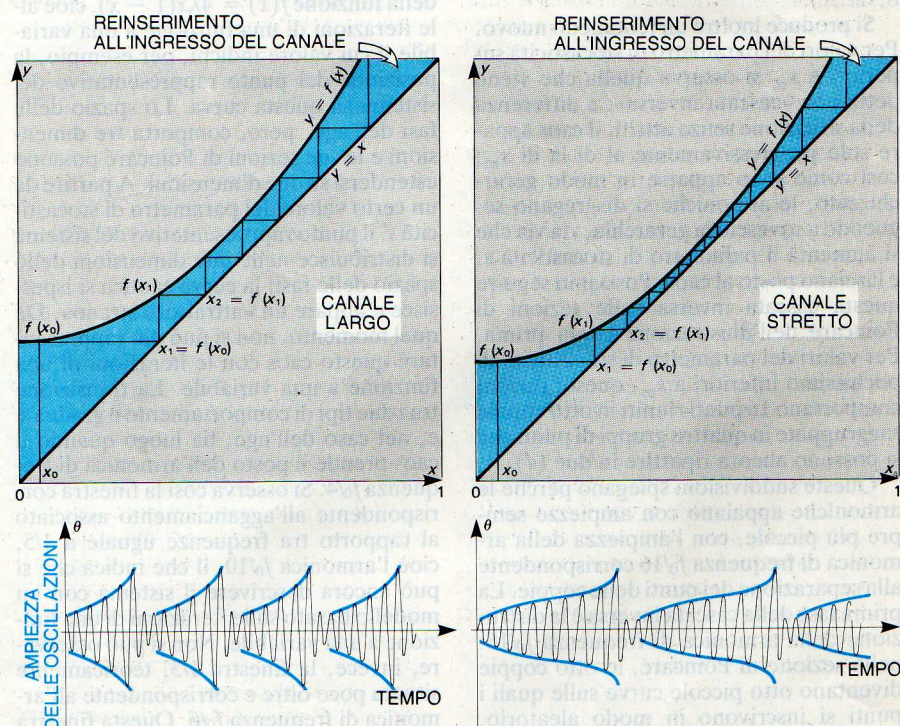
Se si monta l'asse dell'ago su dei rubini, per ridurre gli attriti, è facile osservare i diversi agganciamenti associati ai rapporti Γ tra frequenze uguali a 1/5, 1/4, 1/3, 2/5, 1/2. Vedremo in seguito, infatti, che gli attriti fanno «cadere» il punto rappresentativo del sistema nello spazio delle fasi sui punti ellittici che si trovano al centro delle risonanze. Se si colloca l'ago in campi magnetici tali per cui il parametro di stocasticità s sia uguale a 0,9, come nella sezione di Poincaré dell'illustrazione di pagina 43 (c), a seconda del modo in cui si mette in movimento l'ago si potrà finire nei centri delle cinque risonanze che formano l'agganciamento 1/5 (e l'ago oscilla allora con la frequenza $f_0/5$) oppure al centro della risonanza fondamentale. Con un po' di abilità, si può evidenziare la struttura a bambole russe facendo sì che l'agganciamento 1/3 venga a trovarsi nell'agganciamento 1/5. Si può anche seguire l'inizio della cascata di sdoppiamenti ma, in tutti questi esperimenti, il piccolo contributo degli attriti rende delicata la manipolazione, soprattutto quando si vogliono osservare agganciamenti corrispondenti a piccolissime risonanze nello spazio delle fasi; in tal caso, infatti, il sistema è molto sensibile a ogni perturbazione (per esempio il rumore sperimentale o una modificazione un po' troppo rapida del parametro di stocasticità) e l'ago se ne va rapidamente alla deriva nel mare stocastico che circonda questi isolotti: impazzisce, oscilla a scatti intorno al campo magnetico fisso, si ferma, accelera bruscamente, accompagna il campo magnetico rotante ecc. La stocasticità a grande scala, dunque, è di gran lunga il fenomeno più facilmente osservabile.

Gli attriti influiscono comunque su questo caos; gli episodi di stocasticità a grande scala, infatti, non durano indefinitamente e l'ago riprende in maniera abbastanza improvvisa un moto regolare che corrisponde a un isolotto posto al centro del mare stocastico, vicino al quale è passato il punto rappresentativo prima di cadervi. Tuttavia, quando gli attriti sono deboli, questa erraticità stocastica può durare a volte diverse ore, o anche diversi giorni.

L'importanza degli attriti

Gli attriti non impediscono la comparsa del caos: tutt'al più la ritardano e ne rendono l'evoluzione molto più progressiva. Per studiare le modificazioni apportate dagli attriti, riprendiamo il caso dell'ago collocato in un solo campo magnetico, quello fisso.

Nell'illustrazione di pagina 43 (a) abbiamo rappresentato, nello spazio delle fasi, le traiettorie in assenza di attriti: vi si ritrova la risonanza legata al campo magnetico fisso. In b abbiamo calcolato le traiettorie per le stesse condizioni iniziali, ma questa volta in presenza di attriti: le traiettorie ellittiche sono diventate spirali che terminano nel centro delle risonanze del problema senza attriti. Questo significa che le oscillazioni dell'ago sono smorzate e che il moto d'oscillazione finisce per scomparire. Gli attriti hanno dunque un effetto «di contrazione» nello spazio delle fasi, dato che riportano il punto rappresentativo del sistema al centro di una risonanza, qualunque sia il punto di partenza, purché situato in una certa zona d'attrazione.

Quando l'ago è collocato nei due campi magnetici, in assenza di attriti, lo spazio delle fasi è costituito da numerose risonanze circondate da regioni stocastiche; se gli attriti sono deboli, i centri di risonanza attirano allo stesso modo il punto rappresentativo del sistema. Le risonanze più piccole, però, hanno solo una debole capacità di attrazione, e questa attrazione viene rapidamente superata da quelle delle risonanze più grandi che le circondano; a questo punto le risonanze più piccole non sono più visibili: se si aumentano gli attriti, per esempio ponendo un liquido

L'intermittenza è caratterizzata da uno strano comportamento dell'ago: esso oscilla «regolarmente» intorno al campo magnetico fisso, poi, all'improvviso, l'ampiezza delle oscillazioni aumenta e l'ago compie alcuni movimenti caotici prima di riprendere il suo moto regolare; il ciclo si ripete indefinitamente con periodi di moto quasi regolare la cui durata aumenta progressivamente all'aumentare del parametro di stocasticità. Si può realizzare un modello di questo tipo di comportamento mediante le iterazioni di una funzione a forma di U sull'intervallo [0,1]. I valori consecutivi della funzione possono allora corrispondere, per esempio, all'ampiezza di ciascuna oscillazione dell'ago; su queste curve si vede che la durata della fase quasi regolare (associata alla «traversata» del canale) aumenta quando diminuisce la larghezza del canale tra la curva a U e la bisettrice $y = x$.

viscoso tra l'asse dell'ago e il suo supporto, rimangono osservabili solo le risonanze più grandi. In certi casi (quando gli attriti sono molto rilevanti) le traiettorie terminano solo al centro delle risonanze principali: se si avvia a caso l'ago, esso finirà per allinearsi con il campo magnetico fisso o ruoterà con il campo rotante. In certo qual modo, quindi, esistono solo due tipi di condizioni iniziali.

Se osserviamo il centro della risonanza associata al campo magnetico fisso aumentando ancora il parametro di stocasticità s, appare un fenomeno identico all'esplosione di una risonanza: si produce una cascata di sdoppiamenti di periodo e l'ago, che fino ad allora oscillava intorno al campo magnetico fisso alla frequenza f_0, oscilla inizialmente a una frequenza armonica della fondamentale, $f_0/2$, conservando però contemporaneamente un moto alla frequenza f_0. Un secondo sdoppiamento fa apparire l'armonica $f_0/4$ (*si veda l'illustrazione alle pagine 44 e 45*). Gli attriti facilitano gli esperimenti e si possono osservare fino a quattro sdoppiamenti successivi; l'analogia con i sistemi senza attriti è però abusiva, in quanto, se è vero che essa permette di dare una descrizione semplice dei fenomeni, la cascata di sdoppiamenti che abbiamo appena descritto è differente da quella osservata in assenza di attriti; in particolare, la costante δ, limite del rapporto $(s_n - s_{n-1})/(s_{n+1} - s_n)$, è uguale a 4,66... invece che a 8,72... .

Si produce inoltre un fenomeno nuovo. Per valori del parametro di stocasticità superiori a s_∞ si osserva quella che viene detta una «cascata inversa»: a differenza della situazione senza attriti, il caos appare solo progressivamente al di là di s_∞; così come sono apparse in modo gerarchizzato, le armoniche si disgregano seguendo a rovescio la gerarchia, via via che si aumenta il parametro di stocasticità s, e lasciano posto al caos. Possiamo seguire questa cascata inversa sulle sezioni di Poincaré dell'illustrazione citata prima. Per valori del parametro di stocasticità di pochissimo inferiori a s_∞, queste sezioni comportano 16 punti riuniti in otto coppie raggruppate in quattro gruppi di punti che si possono ancora ripartire in due (*4*).

Queste suddivisioni spiegano perché le armoniche appaiano con ampiezze sempre più piccole, con l'ampiezza della armonica di frequenza $f_0/16$ corrispondente alla separazione dei punti delle coppie. La prima fase della cascata inversa è la distruzione della armonica di frequenza $f_0/16$: nella sezione di Poincaré, le otto coppie diventano otto piccole curve sulle quali i punti si inscrivono in modo aleatorio. Questo caos, però, è molto debole; le otto piccole curve sono ancora ben distinte e lo spettro contiene ancora righe sottili corrispondenti alle combinazioni delle frequenze $f_0/8$, $f_0/4$, $f_0/2$ e f_0. Se si aumenta ancora il parametro di stocasticità s, la stessa situazione si ripresenta per l'armonica $f_0/8$, ma in questo caso il caos diventa un po' più diffuso. Possiamo osservare più facilmente il fenomeno negli altri spettri riportati nell'illustrazione delle pagine 44 e 45: nella sezione di Poincaré (*5*), le otto curve caotiche si sono fuse a due a due, lasciando il posto a quattro curve più marcate. Questo tipo di comportamento è del tutto analogo a quello che si produce per le iterazioni della funzione $f(x) = 4\lambda x(1 - x)$ e si può osservare anche la comparsa delle cosiddette «finestre di rilaminarizzazione»: al centro di questa diffusione del rumore che sostituisce progressivamente l'armonica di frequenza $f_0/8$, si producono agganciamenti associati a rapporti Γ di frequenze uguali prima a 1/5, poi a 1/3, per i quali il sistema torna a essere perfettamente regolare. Lo spettro contiene allora rispettivamente le frequenze $f_0/20$, $f_0/12$ e i loro multipli (*6 e 8 nell'illustrazione delle pagine 44 e 45*).

Gli attrattori strani

Le finestre di rilaminarizzazione sono molto sensibili alle variazioni del parametro di stocasticità, e questo mette in evidenza il carattere deterministico del caos osservato. Dobbiamo anche notare, sulle sezioni di Poincaré, che il caos si sviluppa lungo una curva e non si spande in un nugolo di punti come farebbe, per esempio, un rumore termico. Finché il caos segue una curva a una dimensione, si possono descrivere i comportamenti dell'ago con un modello analogo alle iterazioni della funzione $f(x) = 4\lambda x(1 - x)$, cioè alle iterazioni di una funzione a una variabile il cui valore indichi, per esempio, la posizione del punto rappresentativo del sistema su questa curva. Lo spazio delle fasi dell'ago, però, comporta tre dimensioni e le sue sezioni di Poincaré possono estendersi su due dimensioni. A partire da un certo valore del parametro di stocasticità s, il punto rappresentativo del sistema si distribuisce nelle due dimensioni dello spazio delle fasi: la curva caotica si ispessisce e appare un «attrattore strano». Da quel momento, non si può più rappresentare questo caos con le iterazioni di una funzione a una variabile. La transizione tra i due tipi di comportamento è graduale e, nel caso dell'ago, ha luogo quando il caos prende il posto dell'armonica di frequenza $f_0/4$. Si osserva così la finestra corrispondente all'agganciamento associato al rapporto tra frequenze uguale a 1/5, cioè l'armonica $f_0/10$, il che indica che si può ancora descrivere il sistema con un modello basato sulle iterazioni di una funzione a una variabile. Non si può osservare, invece, la finestra 1/3, teoricamente situata poco oltre e corrispondente all'armonica di frequenza $f_0/6$. Questa finestra dovrebbe essere più grande di tutte quelle già osservate e dunque facile da vedere; in effetti la si osserva quando gli attriti sono rilevanti, ma non si possono osservare le finestre successive, teoricamente ancora più grandi. Quando si aumenta ancora il parametro di stocasticità s, le armoniche subiscono anche in questo caso una distruzione progressiva, ma le sezioni di Poincaré mettono in luce delle «lamelle» nella curva caotica: è comparso un attrattore strano, vera e propria firma del caos. Inizialmente esso corrisponde a oscillazioni irregolari dell'ago intorno al campo magnetico fisso e la sua struttura vista in dettaglio appare costituita da molteplici lamelle leggermente sovrapposte: una struttura quindi molto simile a quella delle bambole russe, caratteristica delle risonanze nel caso senza attriti.

Se si continua ad aumentare il parametro di stocasticità, l'ago finisce con l'accompagnare per qualche giro il campo magnetico rotante nei suoi moti caotici, e l'attrattore strano si espande maggiormente nello spazio delle fasi (*si veda* d *nell'illustrazione di pagina 43*). Questo attrattore strano permane per un intervallo abbastanza ampio di valori del parametro di stocasticità, ma il sistema ritrova ben presto una soluzione regolare, poi ritornà al caos, dopo una nuova cascata di sdoppiamenti. In seguito, il sistema conoscerà un'alternanza di fasi caotiche intervallate da fasi regolari.

L'intermittenza

In occasione di queste alternanze, ogni volta che si passa da una fase regolare a una fase caotica con l'aumento del parametro di stocasticità s, si osservano sia una cascata di sdoppiamenti di periodo sia una cascata inversa. Quando si passa da una fase caotica a una fase regolare, sempre aumentando il parametro di stocasticità s, la transizione avviene con un meccanismo altrettanto interessante: l'intermittenza. Subito prima del ritorno a un movimento regolare, la fase caotica è abbastanza facile da descrivere: l'ago oscilla intorno al campo magnetico fisso e si potrebbe pensare che questa oscillazione sia regolare. In realtà, l'ampiezza dell'oscillazione evolve lentamente nel tempo: prima aumenta molto gradualmente, poi, a partire da un certo punto, si amplifica più nettamente e all'improvviso l'oscillazione diventa caotica (l'ago può accompagnare per qualche giro il campo magnetico rotante, oscillando a scatti, poi rimettersi a oscillare regolarmente, mentre la sua ampiezza evolve a poco a poco ecc.).

Nei moti dell'ago, dunque, si succedono fasi di oscillazioni quasi regolari, intercalate a sprazzi caotici, e la durata di ciascuna di queste fasi è variabile. La durata media delle fasi regolari, tuttavia, varia in funzione del parametro di stocasticità e diventa sempre maggiore al tendere del parametro verso un valore s_r, a partire dal quale l'oscillazione dell'ago è perfettamente regolare. Anche questo fenomeno ha un carattere universale: come nei fenomeni di transizione di fase, la durata media τ delle fasi caotiche varia secondo la legge $\tau = \alpha(s_n - s)^{-0,5}$. Come nel caso degli sdoppiamenti di periodo, Yves Pomeau e P. Manneville, a Gif-sur-Yvette, hanno descritto questo fenomeno servendosi delle iterazioni di una funzione a una variabile: questa funzione ha la forma di

una U su un certo intervallo (per esempio [0,1]), e quando la si itera, il canale che essa forma con la bisettrice (di equazione $y = x$) svolge un ruolo particolare: il punto che rappresenta la posizione dell'n-esimo valore della funzione rimbalza dalla curva alla bisettrice (*si veda l'illustrazione a pagina 47*). Quando il canale è stretto, sono necessarie numerose iterazioni per attraversarlo e, nel corso di questa fase, il valore x_n dell'n-esimo valore della funzione evolve in maniera molto lenta: questo fenomeno corrisponde alla fase regolare dell'intermittenza. Modificare il parametro di stocasticità s equivale a regolare l'altezza della curva a U in rapporto alla bisettrice: il canale si restringe quando il parametro di stocasticità s tende verso s_r. (Per questo valore, la curva è tangente alla bisettrice e i valori successivi della funzione tendono a un limite: il punto di contatto tra la curva e la bisettrice. Il moto dell'ago è allora regolare.) Quando il parametro di stocasticità s è superiore a s_r, la curva a U taglia la bisettrice in due punti, uno dei quali corrisponde alle oscillazioni stabili della fase regolare (l'altro è instabile).

Meccanica classica e fenomeni dissipativi

Se abbiamo deciso di discutere i concetti della meccanica classica non lineare per spiegare l'evoluzione di un sistema dissipativo, che quindi non appartiene a questa classe di problemi, è perché gli sviluppi della meccanica classica hanno svolto senz'altro un ruolo determinante per la comprensione della stocasticità nei sistemi deterministici. In effetti, i progressi nella meccanica classica si sono manifestati in modo abbastanza regolare, e oggi le grandi linee vengono a formare un insieme piuttosto coerente. Il calcolatore ha avuto un ruolo fondamentale nell'elaborazione di concetti nuovi, permettendo di analizzare a fondo i sistemi. Se è vero però che il calcolatore permette di simulare perfettamente un dispositivo meccanico, è altrettanto vero che sono necessari altri mezzi matematici più potenti: la rinormalizzazione, un metodo introdotto da Kenneth Wilson (premio Nobel per la fisica nel 1982) per lo studio delle transizioni di fase, permette di ottenere preziosi risultati quantitativi. Attualmente, D. Escande e F. Doreil dell'École Polytechnique stanno tentando di precisare in modo quantitativo la comparsa della stocasticità a grande scala. La situazione è molto più arretrata nel campo dei problemi dissipativi; si presentano tre grandi scenari per la comparsa del caos (le cascate di sdoppiamenti di periodo introdotte da Feigenbaum, P. Coullet e C. Tresser, il meccanismo di Ruelle-Takens, l'intermittenza descritta da Pomeau e Manneville), ma manca ancora l'unificazione di questi diversi meccanismi; del resto, si scopriranno certamente altri scenari poiché i sistemi dissipativi sono un campo di studi molto più nuovo che non la meccanica.

È comunque confortante constatare la possibilità di illustrare con esperimenti semplicissimi i nuovi concetti apparsi in questi domini: il fatto che il pendolo semplice, o le sue varianti, sia ancora in voga ne sottolinea il carattere fondamentale.

Il moto browniano da Einstein a oggi

La teoria dinamica del moto browniano ci offre un filo conduttore entro la scienza fisica, dalla termodinamica statistica del non equilibrio alla teoria quantistica

di Bernard H. Lavenda

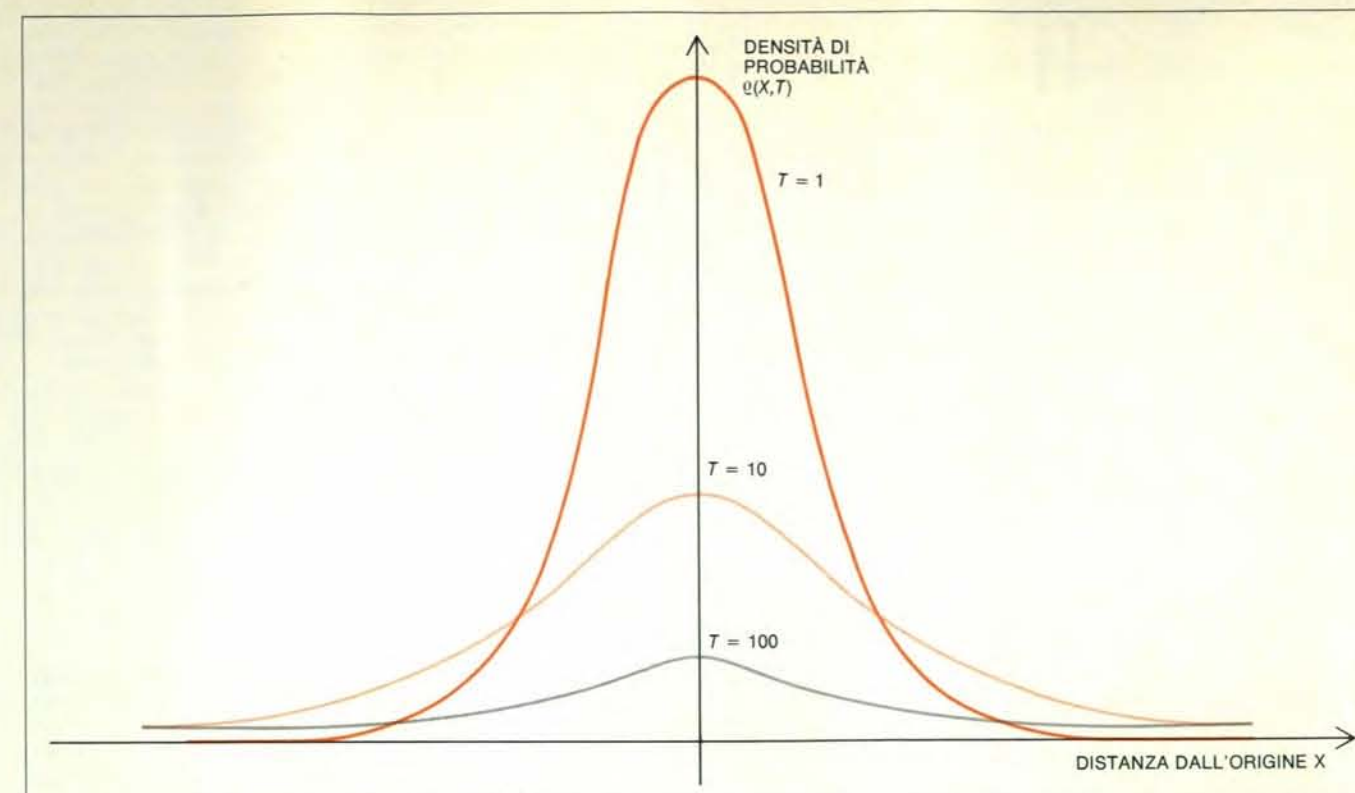
Nel 1827, il botanico inglese Robert Brown fu il primo a osservare al microscopio i minuscoli e rapidi moti irregolari di piccoli granelli di polline sospesi nell'acqua, cioè quello che oggi è conosciuto come «moto browniano». Il suo scopritore fu affascinato dal «rapido moto oscillatorio» dei granelli di polline e dal fatto che «l'inaspettata apparente vitalità di queste "molecole" persiste lungamente dopo la morte della pianta». Le prime spiegazioni attribuivano la causa dei moti irregolari alle correnti termiche convettive all'interno del fluido di sospensione: tuttavia, se così fosse, ci si aspetterebbe che il comportamento di una particella sia correlato con quello delle particelle vicine. Le osservazioni però non confermavano quest'idea, anzi il comportamento di una particella appariva indipendente dal suo passato. Questo dilemma fu indubbiamente causa di grande confusione, in un periodo in cui i pilastri della scienza erano fondati sui principi della meccanica classica: l'osservazione che il moto futuro è indipendente da quello passato, e che il moto è incessante, restò inspiegata per quasi un secolo.

Verso la fine dell'Ottocento si era già messo in evidenza che il moto browniano era tanto più rapido quanto più piccole erano le particelle e quanto più bassa la viscosità del fluido. Anche gli aumenti di temperatura provocavano un aumento della frequenza delle oscillazioni, così che in qualche modo la causa del moto doveva essere imputata ai moti termici delle molecole del mezzo. A quei tempi, la teoria cinetica dei gas era già stata sviluppata grazie al lavoro monumentale di James Clerk Maxwell e Ludwig Boltzmann, durante l'ultima metà del diciannovesimo secolo. Da tale teoria era noto che la temperatura di una sostanza è proporzionale all'energia cinetica media di agitazione delle molecole costituenti il mezzo. Se tale moto di agitazione potesse essere in qualche modo trasferito a molecole sufficientemente grandi da essere osservabili con un microscopio, ciò costituirebbe la prima evidenza diretta della validità della teoria cinetica del calore.

Nel periodo in cui apparve sulla scena Albert Einstein, nel 1905 (lo stesso anno in cui formulò la teoria della relatività ristretta, e in cui spiegò l'effetto fotoelettrico introducendo il concetto di «fotone»), era disponibile una notevole mole di dati sperimentali sul moto browniano. Ciononostante, Einstein era alla ricerca di fatti che comprovassero l'esistenza di atomi di dimensioni definite e non conosceva quel fenomeno; i suoi sforzi lo portarono a prevederne l'esistenza su basi puramente teoriche, e a fornirne la prima teoria quantitativa.

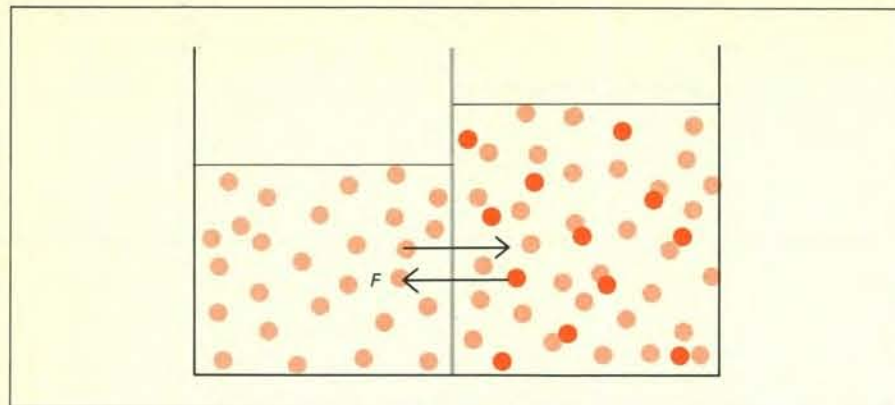
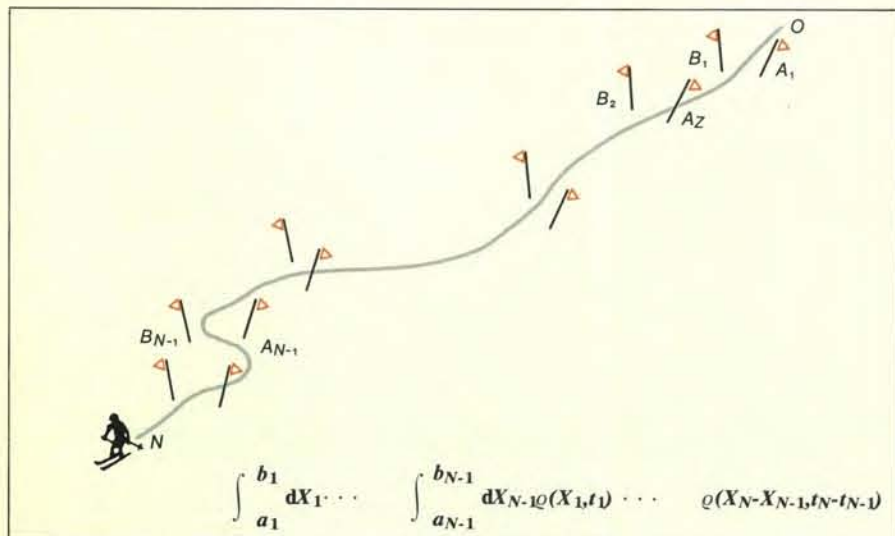
Consideriamo il moto di una particella libera, sulla quale, cioè, non agisce alcun campo di forze esterne: per determinare il moto della particella non basterebbe conoscere gli impulsi che essa riceve in un dato intervallo di tempo, ma occorrerebbe sapere anche la sua velocità iniziale. Per Einstein la conoscenza della velocità iniziale della particella browniana, rispetto a un qualsiasi intervallo di tempo di osservazione, rappresentava un dato insignificante in confronto al numero di urti che la particella riceve nello stesso intervallo di tempo. (In effetti la particella browniana subisce circa 10^{21} collisioni al secondo; quindi l'assunzione di Einstein è ampiamente giustificata.)

Facendo l'ulteriore assunzione di una distribuzione casuale delle posizioni molecolari e prendendo in considerazione intervalli di tempo lunghi rispetto all'intervallo di tempo medio fra due collisioni molecolari successive, Einstein giunse alla formulazione di un'equazione di diffusione, analoga a quella che descrive la conduzione termica, la cui soluzione fornisce la densità di probabilità che una particella browniana occupi una data posizione a un dato istante. Se consideriamo tale funzione nel caso di una particella che si trovi inizialmente nell'origine e diffonda in una sola direzione, la densità di probabilità si comporta in modo simile a una goccia di inchiostro che diffonde nel tempo in un bicchiere d'acqua. In altre parole, potremmo immaginare la densità di probabilità come la distribuzione, lungo una direzione dello spazio, di una sostanza estranea introdotta in un mezzo omogeneo, e l'evoluzione di questa densità di probabilità nel tempo come la diffusione della stessa sostanza. Per avvicinarci di più al modello reale, la posizione di un punto rappresentativo della sostanza a ogni dato istante dovrebbe essere descritta da una funzione casuale, associata al moto di ciascuna molecola della sostanza. La traiettoria deve avere necessariamente una forma estremamente complicata e discontinua, il che determina una caratteristica di vitale importanza del moto browniano: la velocità istantanea di un punto che descrive la traiettoria del processo non è definibile!

Comunque, in un intervallo di tempo finito, si può ottenere uno spostamento finito per il fatto che la velocità del punto rappresentativo inverte il suo segno con frequenza infinita, mentre il punto si muove in entrambe le direzioni. Così, dall'osservazione di un grande numero di processi casuali, otteniamo un gran numero di punti rappresentativi che si spostano in maniera erratica, o casuale, come in un movimento a «zig-zag», a causa delle interazioni con le particelle del mezzo. La pendenza di ogni tratto di cammino libero (lo «zig») non è necessariamente uguale a quella di un altro (lo «zag») e, con l'aumentare della frequenza delle interazioni, il «cammino libero medio» della particella diminuisce. Al limite dell'idealizzazione matematica, in cui il cammino libero medio tende a zero, possiamo dire che il vettore spostamento di una particella browniana non è differenziabile in alcun punto; non possiamo cioè definire una velocità per il processo. Conseguentemente, tutto quello che possiamo fare è parlare di una densità di probabilità, che equivale ad avere una densità di punti costituenti una sorta di gas che diffonde. Naturalmente queste particelle di gas non possono essere né create né distrutte durante il loro moto e ciò equivale a dire che la probabilità si conserva; il valore della probabilità in una certa regione corrisponde al numero di punti che si trovano in quella regione oppure, equivalentemente, corrisponde al tempo che in media ogni «punto browniano» trascorre in quella regione.

La densità normale o gaussiana rappresenta la quantità $\varrho(X,t) = [4\pi Dt]^{-1/2} \times \exp(-X^2/4Dt)$. Possiamo immaginare un gran numero di particelle simili, affollate nelle immediate vicinanze di $X = 0$ all'istante $t = 0$, che vengono lasciate a se stesse da $t = 0$ in poi. Dopo un tempo t, si stabilisce spontaneamente una distribuzione di particelle tale che il relativo numero di particelle comprese fra X e $X + dX$ è : $\varrho(X,t)dX$. Oppure, alternativamente, possiamo considerare come nostro sistema, non un gran numero di particelle simili fra loro ma, piuttosto, una particella singola. Allora $\varrho(X,t)dX$ denota la probabilità che la particella si sia spostata, nel tempo t, in una regione compresa fra X e $X + dX$. Sulla base di questa formula, Einstein calcolò che l'allontanamento medio da $X = 0$ doveva essere $\sqrt{2Dt}$ al tempo t, in cui D è il coefficiente di diffusione. Einstein concluse così che il cammino descritto in media da una particella non è proporzionale al tempo, ma alla radice quadrata del tempo. Ciò deriva dal fatto che gli spostamenti descritti, per esempio durante due intervalli di tempo unitari, «non vanno sempre sommati fra loro, ma altrettanto frequentemente vanno sottratti».

Anche se questa descrizione costituisce un'astrazione limite dei processi che avvengono realmente in natura, è la sola che rende possibile allo stesso tempo l'introduzione di concetti probabilistici e l'interruzione della trasmissione di informazione tra passato e futuro. In altre parole, per descrivere il moto browniano, occorre simulare un processo senza «memoria», utilizzando quella che i matematici chiamano proprietà markoviana. Del resto, qualsiasi processo casuale, anche con memoria, può sempre essere decomposto in processi più elementari che godano della proprietà markoviana (ne troverete una prova vivente nella maggior parte delle istituzioni burocratiche esistenti).

Nella teoria di Einstein compare un solo parametro caratteristico - il coefficiente di diffusione - e Einstein derivò una formula in cui esso era espresso in termini del numero di Avogadro e di altre grandezze fisiche che potevano essere misurate in laboratorio. Einstein suppose che la diffusione delle particelle sospese nel liquido fosse governata da una condizione di «equilibrio dinamico» fra la forza osmotica che tende a sospingere le particelle dalle regioni a alta concentrazione alle regioni a bassa concentrazione, e una forza viscosa che tende a ritardare il moto delle particelle. La forza viscosa è proporzionale alla velocità della particella anziché alla sua accelerazione, poiché l'accelerazione iniziale subisce un rapido smorzamento in un mezzo viscoso. Einstein voleva evitare il ricorso alla nozione di velocità ben definita della particella e la novità del suo trattamento consiste nel tentativo di descrivere il moto delle particelle con ragionamenti probabilistici. Infatti l'analisi di Einstein prelude allo sviluppo della teoria matematica dei processi stocastici e l'eleganza del suo procedimento consiste nel fatto che la velocità introdotta con la forza viscosa è puramente virtuale! La formula di Einstein per il coefficiente di diffusione si applica anche quando non è definita alcuna velocità e anche quando vi è una sola particella browniana, per cui non è possibile definire la concentrazione! Allora, semplicemente invertendo la formula di Einstein, era possibile ottenere il numero di Avogadro dalla misura del coefficiente di diffusione di una sospensione colloidale di particelle sferiche di raggio approssimativamente uniforme. Considerando il numero di ipotesi introdotte nella derivazione della formula di Einstein, è veramente notevole che J. Perrin ottenesse un risultato sperimentale in accordo entro il 19 per cento col valore effettivo del numero di Avogadro, dedotto per altre vie.

Tutto ciò che abbiamo detto riguardo alla teoria di Einstein non le rende completa giustizia. Grazie agli sforzi di Einstein venne alla luce che la meccanica statistica era una teoria con implicazioni sperimentali che non potevano essere spiegate dalla termodinamica classica. La teoria delle fluttuazioni inaugurata dalla teoria di Einstein ha dato frutti che solo di recente cominciano a essere presi in considerazione dalle scienze fisiche e matematiche. Durante gli ultimi due decenni è stata sviluppata un'intensa ricerca, che nella scienza matematica va sotto il nome di «studio dei processi stocastici di diffusione». Sono state trovate immediate applicazioni nelle teorie di ottimizzazione dei controlli e del filtraggio dei segnali. Ma applicazioni ancora più ampie alle scienze fisiche e chimiche hanno portato a dimostrare che la teoria del moto browniano può sia costituire le fondamenta

Un esperimento sull'osmosi illustra la situazione di equilibrio «dinamico». Una membrana semipermeabile, permeabile alle molecole di solvente *(in colore chiaro)* e impermeabile alle molecole di soluto *(in colore più intenso)*, separa due compartimenti: uno contiene molecole di solvente soltanto e l'altro contiene sia molecole di solvente che di soluto. Il gradiente di concentrazione delle molecole di soluto dà luogo a una forza osmotica che agisce nel senso di aumentare il flusso di molecole di solvente nella direzione da sinistra verso destra. Nello stato di equilibrio dinamico, la forza osmotica F_O è bilanciata da una forza uguale ed opposta F_V, che è la forza viscosa che agisce in modo da ritardare il moto delle molecole di solvente. La differenza di livello fra i due compartimenti è uguale alla pressione osmotica del solvente.

Lo slalom (ottenuto con una simulazione al calcolatore) illustra il modo di Wiener di trattare il moto browniano come «somma su tutti i cammini». Sapendo che lo sciatore parte da *O* e arriva a *N*, la probabilità che esso passi attraverso ciascuna delle porte, negli istanti di tempo associati, è data dall'integrale multiplo riportato in basso a destra nell'illustrazione, in cui ϱ è la densità di probabilità visualizzata nell'illustrazione della pagina precedente (cioè la distribuzione trovata da Einstein). È piuttosto intuitivo che, se aumentassimo il numero di ostacoli e se avvicinassimo sempre più i paletti delle porte fra loro, potremmo tracciare il cammino dello sciatore sempre con maggior precisione. Andando al limite, otterremo una misura nello «spazio dei cammini» nota come misura condizionale di Wiener, condizionata dalla nostra conoscenza del fatto che lo sciatore è partito da *O* ed è arrivato a *N*.

della termodinamica statistica dei processi di non equilibrio sia (anche se meno rigorosamente, dal punto di vista matematico) fornire un completamento alla formulazione della meccanica quantistica introdotta da Richard P. Feynman mediante i cosiddetti «integrali di cammino». Nel seguito, vogliamo presentare alcune delle nuove e eccitanti prospettive di ricerca chimico-fisica che la teoria del moto browniano ha recentemente aperto.

Sebbene molti studi sulla natura fisico-matematica del moto browniano fossero stati effettuati prima del 1923, in particolare da Einstein, M. Smoluchowski, P. Langevin, J. Perrin e altri, la formulazione matematica completa di quella che è oggi nota come teoria del moto browniano fu presentata da N. Wiener, nel suo ormai famoso lavoro sugli spazi differenziali. Per questo nella letteratura matematica spesso ci si riferisce al moto browniano col nome di «processo Wiener». In quegli anni, Wiener sviluppò un'interpretazione del moto browniano come «somma su tutti i cammini», che descriveremo qualitativamente. Consideriamo una particella che subisca una serie di spostamenti, tali che l'entità e la direzione di ciascuno di essi sia indipendente dagli spostamenti precedenti. Ora, la probabilità che lo spostamento della particella browniana avvenga fra i due punti a_1 e b_1 è determinata da una funzione di distribuzione che, se la particella ha un moto simile al cammino di un ubriaco, risulta essere proprio la soluzione dell'equazione di diffusione derivata da Einstein. Il problema è determinare qual è la probabilità che, dopo *n* spostamenti, la particella venga a trovarsi nella zona fra a_n e b_n. Così, la probabilità dipende da un gran numero di altre grandezze aventi distribuzioni preassegnate entro un intervallo di valori. Il problema che ci troviamo di fronte è analogo a quello di uno sciatore che debba affrontare un percorso di slalom *(si veda l'illustrazione in questa pagina in basso)*. La probabilità che l'atleta passi attraverso una porta-ostacolo è data dal prodotto della larghezza della porta per la densità di probabilità, fattore peso che tiene conto sia della mobilità dello sciatore, sia della distanza fra due porte successive, sia ancora del tempo impiegato dallo sciatore per passare da una porta alla successiva. Se ora osserviamo l'atleta passare attraverso una data porta in un certo istante, potremmo desiderare di conoscere qual è la probabilità che egli passi attraverso un successivo ostacolo dopo un certo lasso di tempo. Se tale intervallo è grande, il fatto che lo sciatore passi o meno attraverso la porta non dovrebbe dipendere dal fatto che noi lo abbiamo visto passare per un'altra porta in un istante precedente. La probabilità totale sarà allora il prodotto delle probabilità individuali e possiamo immaginare che, aumentando il numero di osservazioni sullo sciatore (cioè aumentando il numero degli ostacoli sul suo percorso) e rendendo sempre più piccola la larghezza di ciascuna porta, saremmo capaci di localizzare la traiettoria dello sciatore con sempre maggiore precisione.

La difficoltà risiede nell'andare al limite di osservazioni infinitamente frequenti, e tale è appunto il problema risolto da Wiener. L'indipendenza statistica tra gli eventi può essere rigorosamente mantenuta anche nel caso limite di piccoli intervalli di tempo?

Al limite di intervalli infinitamente piccoli e di ostacoli infinitamente stretti, otteniamo quella che i matematici conoscono come «misura» di Wiener. La misura è proprio il numero che otteniamo facendo il prodotto delle varie probabilità individuali per ogni singolo evento. In più, quando non è richiesta la conoscenza del passaggio dello sciatore attraverso una qualsiasi particolare porta-ostacolo, dobbiamo sommare le probabilità su tutti i punti attraverso i quali lo sciatore potrebbe essere passato. Possiamo ora collegare il problema degli ostacoli dello slalom alla teoria delle misure, sia nella teoria classica, sia in quella quantistica.

In entrambe queste teorie il concetto di probabilità di ottenere un dato risultato in un esperimento è fondamentale. Le nozioni probabilistiche entrano nella fisica classica a causa della impossibilità di conoscere posizione e velocità di particelle in un sistema macroscopico nel quale la popolazione sia dello stesso ordine di grandezza del numero di Avogadro (circa 10^{23}). Anche se potessimo seguire ciascuna particella singolarmente, le nostre misure eseguite con strumenti macroscopici, pur non introducendo perturbazioni nel moto, potrebbero non rispecchiare il comportamento medio delle singole particelle. In meccanica quantistica, le considerazioni probabilistiche entrano per un'altra ragione: in questo caso abbiamo a che fare con particelle di dimensioni estremamente ridotte (per esempio elettroni), così che qualunque apparecchio di misura usato per rilevarne la posizione introduce una perturbazione inevitabile; le particelle che interagiscono durante la misura, per esempio i fotoni, hanno dimensioni dello stesso ordine di grandezza degli oggetti che vogliamo misurare! Classicamente potremmo chiederci qual è la probabilità P_{ab} che la misura *A* dia il risultato *a* e che, nello stesso tempo, la misura *B* dia il risultato *b*. In maniera simile, poniamo che P_{bc} sia la probabilità che la misura *B* dia il risultato *b* mentre la misura *C* dia il risultato *c*; supponiamo inoltre che P_{abc} sia la probabilità relativa al verificarsi di tutti e tre i risultati: cioè *A* dà *a*, *B* dà *b* e *C* dà *c*.

Allora, se gli eventi fra *b* e *c* sono indipendenti da quelli fra *a* e *b* (assumendo che le misure *A*, *B*, *C* vengano eseguite in successione temporale con lo stesso ordine), la probabilità è semplicemente il prodotto:

$$P_{abc} = P_{ab} \times P_{bc}.$$

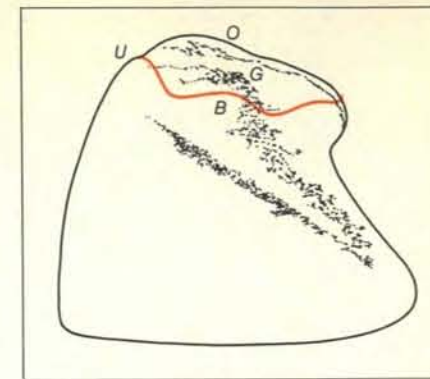
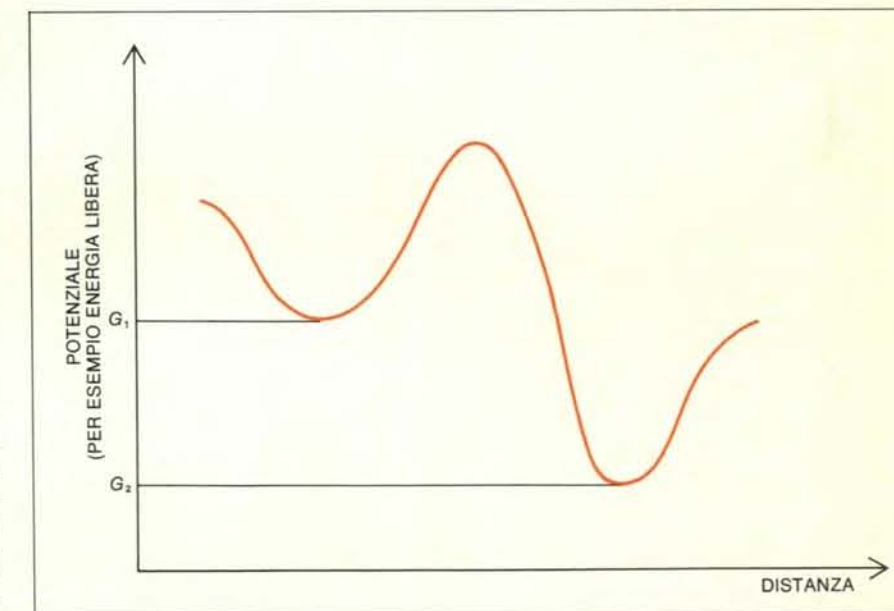
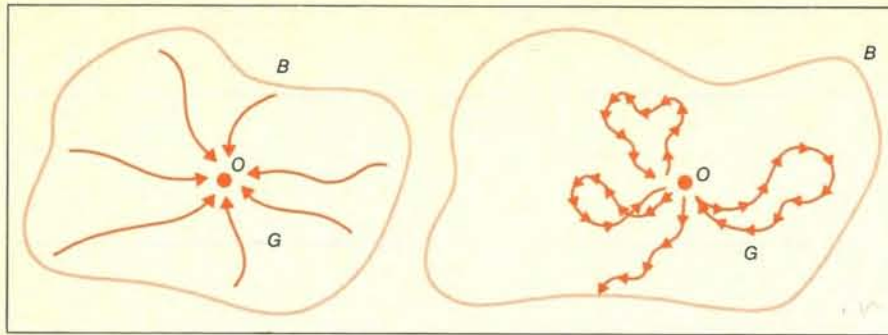
Supponiamo ora di non eseguire la misura *B*; la probabilità che *A* dia *a* e che *C* dia *c* è proprio:

$$P_{ac} = \text{somma su tutti i } b\ (P_{abc}),$$

questo poiché la grandezza *b* deve necessariamente assumere qualche valore fra le misure *A* e *C*.

Classicamente questa seconda equazione è corretta, mentre è stato dimostrato che in meccanica quantistica essa è in generale errata. Perché? Abbiamo dovuto ipotizzare che, passando da *a* a *c*, *B* abbia assunto qualche valore definito *b*. Ora, qualsiasi tentativo di misura disturberebbe il sistema in modo tale da rendere l'equazione errata per la meccanica quantistica. La perturbazione introdotta dall'apparecchio di misura sull'oggetto che deve essere misurato sarebbe sufficiente per trasformare il nostro oggetto in un sistema completamente nuovo dopo la misura! Che la seconda equazione non possa essere vera in meccanica quantistica fu affermato chiaramente per la prima volta da Werner Heisenberg col suo famoso principio di indeterminazione.

Classicamente è possibile caratterizzare un cammino con una successione di misure a istanti successivi che ci forniscano una successione di punti. Se eseguiamo un numero sufficientemente alto di misure, possiamo collegare i punti e definire una traiettoria. Sia $P_{r_1 r_2 r_3 \ldots}$ la probabilità per tale cammino. Qui r_1, r_2,... sostituiscono i risultati sperimentali *a*, *b*, *c*,... Se desideriamo conoscere la probabilità che r_i sia fra a_i e b_i, ecc., dobbiamo sommare (più correttamente integrare) su tutti i possibili valori intermedi. Se usassimo la soluzione dell'equazione di diffusione di Einstein come densità di probabilità, il risultato ottenuto sarebbe la misura di Wiener.

Ora, in meccanica quantistica non è possibile seguire un cammino di una particella, perché ogni volta che misuriamo la sua posizione ne disturbiamo il percorso tanto da convertirlo in un altro processo. Feynman fu il primo a capire che se si fossero sostituite le probabilità con le «ampiezze di probabilità», allora tutte le regole della probabilità classica sarebbero state valide per la meccanica quantistica. Secondo l'interpretazione ortodossa della meccanica quantistica, le ampiezze di probabilità sono grandezze complesse i cui moduli quadrati forniscono le densità di probabilità. Poiché le ampiezze di probabilità stesse non costituiscono grandezze fisicamente osservabili, è lecito intenderle quale strumento probabilistico per la descrizione dei cammini delle particelle. Tutto ciò che cambia, passando dalla visione classica di Wiener di una somma su tutti i cammini all'interpretazione di Feynman, sta nell'introduzione di una misura «complessa». Non abbiamo alcun motivo per opporci a tale assunzione perché i cammini non sono in alcun modo

Una riproduzione della superficie entropica che venne presentata a Gibbs da Maxwell nel 1875. La funzione entropia può essere rappresentata graficamente con una superficie in un sistema multidimensionale di coordinate generalizzate, con assi che indicano i valori della differenza fra le variabili termodinamiche estensive generalizzate e quelle rispettivamente corrispondenti ai valori di equilibrio. Lo stato di massima entropia *O* corrisponde allo stato di equilibrio. Considerando qualunque dominio arbitrario *G*, con frontiera *B*, il sistema, quasi certamente, farà la sua uscita allo stato *U*, che ha la massima entropia rispetto a tutti gli altri stati della frontiera, a causa delle fluttuazioni termiche che guideranno il sistema al di fuori di qualsiasi dominio circoscritto, contenente lo stato di equilibrio, nel limite in cui l'intensità del rumore termico tende a zero. Nella stessa condizione, il sistema non visiterà quegli stati le cui entropie siano minori di quella massima sulla frontiera *B*, a patto che esista uno stato di massima entropia in *B*.

Il diagramma raffigura una buca di potenziale bistabile: un campo di forze con due punti di stabilità locale. Se prendiamo in considerazione le fluttuazioni, una particella ha qualche possibilità di superare la barriera di potenziale, e così le stabilità relative delle due buche possono essere messe a confronto fra loro. Questa descrizione viene spesso usata come un modello per le reazioni chimiche, e in tal caso la distanza denota una coordinata di reazione. Le collisioni molecolari possono eccitare una molecola tanto da superare la barriera, così che si verifichino la rottura del legame molecolare e la dissociazione della molecola. Le collisioni casuali sono descritte da un «rumore bianco» la cui intensità determina la temperatura dei reagenti. Una volta superata la barriera di potenziale, c'è la possibilità che si vengano a formare nuovi e diversi legami molecolari, e che di conseguenza il sistema vada a cadere nello stato di equilibrio più stabile, restandovi in quiete (a meno delle fluttuazioni termiche). Il tasso con cui le collisioni molecolari spingono le molecole al di là della soglia della buca di potenziale è determinato dalla cinetica della reazione.

A sinistra le linee di flusso per il caso in cui lo stato stazionario O, nel dominio G, è stabile. Le fluttuazioni termiche casuali possono essere immaginate come causa di una lenta diffusione del sistema in un campo di flusso deterministico. Poiché lo stato stazionario è stabile, il sistema deve diffondere «contro il flusso» per raggiungere il contorno B. A causa delle fluttuazioni termiche casuali, il sistema può passare da uno stato stazionario stabile ad un altro, ma esso deve necessariamente andare contro il flusso, almeno per una parte della traiettoria. A destra, il moto del sistema, precedente alla sua uscita dal contorno B, a partire da un intorno dello stato stazionario stabile O appartenente a un dominio limitato G, ha la seguente natura: il sistema è attratto dallo stato stazionario, qualunque sia lo stato in G in cui esso si trovi, e il moto verso lo stato stazionario è descritto da un'equazione deterministica cinetica. Giunto nelle vicinanze di O, il sistema può effettuare escursioni in quegli stati che sono più distanti, per poi tornare indietro attratto dallo stato stazionario. Se accade che il sistema raggiunge il contorno, allora il suo moto, in precedenza, non sarà lento e continuo, ma piuttosto avrà l'aspetto di un moto a balzi in cui percorre una distanza definita in un tempo definito: altrimenti verrebbe risucchiato verso lo stato stazionario.

osservabili. Il rigore matematico di tale formulazione è ancora in fase di perfezionamento poiché insorgono problemi di convergenza, e in più non è stata individuata la provenienza del rumore che impedisce una trattazione deterministica della meccanica quantistica.

Wiener mise in relazione la densità di probabilità con il processo di diffusione di una particella «libera», mentre Feynman diede, in forma assiomatica, regole per il calcolo della densità di ampiezza di probabilità, dimostrando a posteriori che essa soddisfa l'equazione di diffusione più notevole della meccanica quantistica: l'equazione di Schrödinger. Ha senso allora chiedersi a quale processo fisico queste regole formali corrisponda-no. Per rispondere a questa domanda dobbiamo introdurre una linea di ricerca parallela allo studio delle equazioni di diffusione: la teoria delle equazioni differenziali stocastiche. P. Langevin può essere considerato il fondatore di tale formalismo che consente una rappresentazione semplice (sebbene matematicamente complessa) del processo fisico chiamato moto browniano.

Caratteristica di questa interpretazione è il fatto che le equazioni del moto utilizzate nell'analisi dei sistemi fisici sono una riformulazione della legge del moto di Newton: $F = ma$, cioè forza = massa × accelerazione, con l'eventuale inclusione delle forze di attrito, o dissipative. Sebbene l'applicazione diretta di $F = ma$ conduca a equazioni differenziali del secondo ordine, è sempre possibile, mediante l'introduzione di variabili aggiuntive, trasformare queste equazioni in un sistema di equazioni differenziali accoppiate del primo ordine (ma sovente non lineari), aventi la forma:

$$\frac{dx(t)}{dt} = b[x(t)]$$

dove $x(t)$ è in generale un vettore di dimensione $n>1$, detto «stato del sistema» al tempo t. Se si vogliono prendere in considerazione anche fluttuazioni del sistema simili al moto perpetuo delle particelle browniane, mantenute in tale stato dalle collisioni casuali con le molecole più leggere del fluido circostante, occorre aggiungere, come di consueto, una forza casuale o fluttuante al membro di destra di quest'ultima equazione; ciò provoca la conversione dell'equazione differenziale deterministica in un'altra, nota come equazione differenziale stocastica (o equazione di Langevin), avente la forma:

$$\frac{dx(t)}{dt} = b[x(t)] + f(t).$$

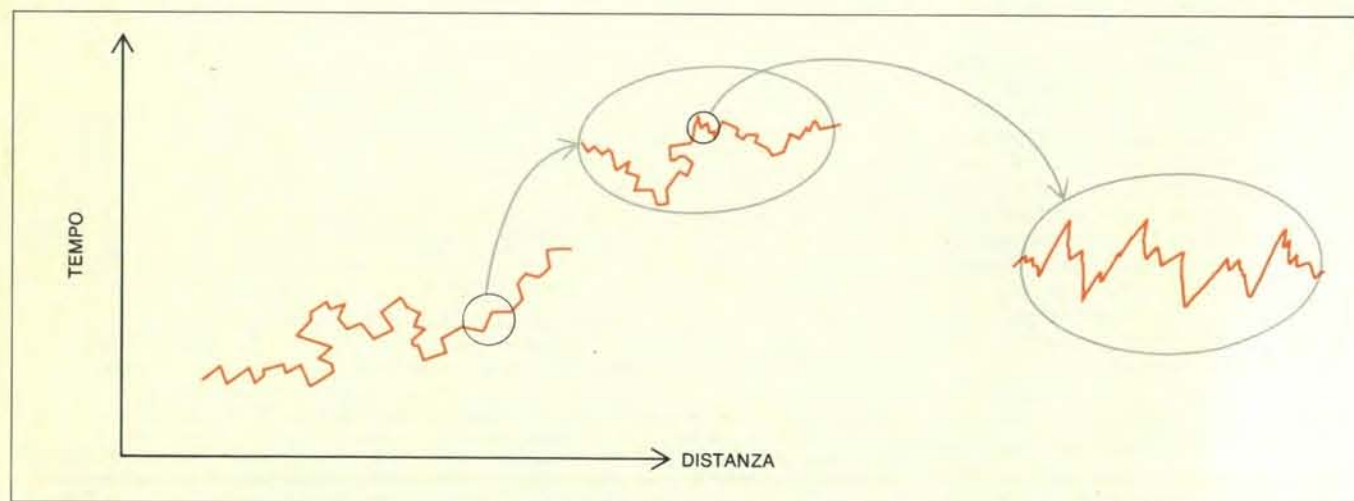
Le equazioni differenziali del tipo dell'equazione di Langevin connettono due mondi separati: il mondo macroscopico rappresentato dal «vettore di deriva» b, e il mondo microscopico, rappresentato dalla forza fluttuante f. Nella formulazione originale di Langevin, $x(t)$ rappresentava la quantità di moto della particella browniana e $b[x(t)]$ l'attrito dinamico agente sulla particella; la parte fluttuante $f(t)$ caratterizzava il moto browniano. Langevin arguì che tale decomposizione doveva essere valida in quanto il moto ha luogo su due scale di tempo largamente separate fra loro: una corta, secondo la quale varia rapidamente la forza fluttuante $f(t)$ (ricordiamo che una particella subisce normalmente circa 10^{21} collisioni al secondo) e una lunga, rispetto alla quale si manifestano gli effetti dell'attrito dinamico.

Risolvere un'equazione differenziale stocastica del tipo di quella di Langevin non è come risolvere un'equazione differenziale ordinaria: infatti l'equazione di Langevin mette in gioco una forza fluttuante $f(t)$ che ha proprietà definite solo statisticamente. In assenza di una conoscenza specifica delle forze casuali, si assume comunemente che $f(t)$ sia un cosiddetto processo casuale gaussiano di «rumore bianco»: il termine «rumore bianco» è usato per analogia con la luce «bianca», poiché in entrambi i casi lo spettro è costante, ovvero di ampiezza indipendente dalla frequenza. Si deve tuttavia considerare che un vero rumore bianco non può esistere nel mondo reale: qualsiasi rumore osservabile, infatti, indipendentemente da quanto risulti «piatto» il suo spettro alle basse frequenze, si annulla regolarmente alle alte frequenze. Il modello del rumore bianco, quindi, porta inevitabilmente alla cosiddetta «catastrofe ultravioletta», conseguenza questa che vanificò i tentativi di trattare lo spettro del «corpo nero» sulla base del principio di equipartizione dell'energia. Perciò si tratta di una idealizzazione matematica che può soltanto approssimare la realtà.

In effetti esistono altre forme di rumore; per esempio il rumore granulare creato per emissione spontanea dal catodo di un elettrone che, raggiungendo l'anodo, produce una corrente. Analogamente l'emissione di altri elettroni è descritta da una distribuzione casuale di tempi di emissione che statisticamente ha l'aspetto di una distribuzione di Poisson. Comunque, se applichiamo la legge dei grandi numeri, sappiamo che al limite di molti eventi indipendenti le distribuzioni statistiche tendono sempre a una distribuzione gaussiana, e ciò ci riconduce al rumore bianco che è sempre stazionario nel tempo e ha uno spettro uniforme.

A questo punto viene da chiedersi: quale delle due descrizioni, deterministica (equazioni differenziali del primo ordine) o statistica (equazioni differenziali stocastiche di Langevin), si avvicina di più a una descrizione realistica di ciò che accade nei sistemi dinamici? Prendiamo in considerazione, per esempio, la termodinamica classica: essa prevede che le fluttuazioni relative di una variabile termodinamica estensiva siano proporzionali all'inverso della radice quadrata del numero delle particelle. Al limite termodinamico, in cui il numero delle particelle e il volume tendono all'infinito in modo tale che il loro rapporto rimanga costante, le fluttuazioni relative tendono a zero, e la distribuzione si raccoglie sempre più attorno al valore atteso, che è quello che la termodinamica prevede essere il valore sperimentale. Ma noi sappiamo che si verificano sempre piccole deviazioni dalle equazioni di stato termodinamiche: se due sistemi sono preparati in modo identico, non è detto che successivamente si comporteranno esattamente allo stesso modo.

È questo che intendiamo quando sosteniamo la necessità di prendere in considerazione le fluttuazioni. Inoltre, in piccole regioni di spazio, oppure in prossimità di un punto critico, le proprietà di un sistema fluttuano largamente intorno ai valori previsti dalle equazioni di stato deterministiche, e il comportamento macroscopico futuro è determinato dalle fluttuazioni iniziali. Consideriamo inoltre il caso di una buca di potenziale bistabile, del tipo di quelle comunemente usate per descrivere le reazioni chimiche *(si veda l'illustrazione a pagina 000)*. La termodinamica classica prevede che il sistema tenderà a trovarsi sempre in corrispondenza del minimo più basso (G_2) perché esso corrisponde all'energia libera più bassa; un sistema che si trovasse però nel minimo più alto (G_1) sarebbe destinato a rimanervi, in assenza di fluttuazioni: quindi la previsione della termodinamica classica non può essere del tutto legittima. In altri termini non esiste per il sistema alcun modo per accorgersi della presenza di un minimo dell'energia libera termodinamicamente più stabile.

È ben noto che le fluttuazioni giocano un ruolo cruciale nel caso di sistemi fisici che diventano instabili (in modo appros-

I cammini tipici di una particella browniana sono fortemente irregolari se osservati su scala molto fine, come mostrato qui. Così, è possibile definire una velocità media, ma non esiste una velocità istantanea. Questo significa che i cammini del moto browniano non sono differenziabili.

Mentre l'entropia del sistema tende ad aumentare fino a raggiungere il valore massimo all'equilibrio, l'entropia generata dalle correlazioni statistiche fra stati di non equilibrio tende, in media, a diminuire col trascorrere del tempo. L'entropia addizionale, prodotta dalle correlazioni statistiche, distrugge la proprietà termodinamica di additività, che torna a valere solo per tempi lunghi, ove le correlazioni statistiche abbiano avuto abbastanza tempo per esaurirsi.

Schematizzazione delle fluttuazioni intorno al cammino medio per la regressione del sistema fino all'equilibrio. Poiché le medie e i valori più probabili di una distribuzione gaussiana sono identici, il cammino medio coinciderà con il cammino più probabile per la regressione di una fluttuazione solo nel caso che le fluttuazioni siano gaussiane. Per fluttuazioni non-gaussiane, solo nel limite di intensità di rumore termico molto piccola, le distribuzioni delle variabili termodinamiche fluttuanti diverranno così nette che il comportamento medio e quello più probabile torneranno a coincidere.

simativo possiamo dire che un sistema dinamico è stabile se è insensibile a piccole perturbazioni). Comunque le nostre equazioni, sia quella deterministica, sia quella di Langevin, prevedono un comportamento qualitativamente differente. Assumiamo che esista uno stato deterministicamente stabile, isolato e stazionario, determinato dalla scomparsa della deriva: $b(O) = 0$ in cui con O si simboleggia lo stato stazionario. Questo significa che il sistema non sta «evolvendo». Allora ambedue le equazioni saranno concordi nel prevedere che il sistema tenderà inizialmente ad agire in modo da eliminare ogni eventuale variazione del suo stato causata da perturbazioni, ristabilendo così lo stato stazionario, O.

Ambedue le equazioni sono concordi nel dire che il sistema «scivolerà» lungo una curva che sia soluzione del sistema deterministico. Ora l'equazione differenziale del primo ordine prevede che il sistema si avvicinerà asintoticamente allo stato stazionario, e null'altro. Al contrario l'equazione di Langevin prevede che il sistema si avvicinerà a un piccolo intorno di O, nel quale trascorrerà la maggior parte del suo tempo. Ma, tenendo conto della forza casuale, il sistema avrà sempre una probabilità di saltare in un altro stato in G. Il moto non è continuo; consiste invece di salti discontinui, poiché il sistema avverte sempre l'attrazione dello stato stazionario. Presto o tardi, il sistema raggiungerà la frontiera B con una probabilità prossima a 1. Certamente il tempo che impiegherà per raggiungere tale frontiera sarà più lungo di quello impiegato in tentativi infruttuosi di raggiungere B. Comunque l'equazione di Langevin dice che il sistema abbandonerà quasi certamente qualsiasi dominio circoscritto in un tempo finito. Il sistema è sempre instabile, indipendente dalla grandezza della forza fluttuante, posto che tale forza agisca anche nello stato stazionario che dall'equazione differenziale ordinaria del primo ordine risulta essere globalmente stabile (la forza fluttuante è in media zero).

Una simulazione col metodo di Monte Carlo di un sistema dinamico bidimensionale lineare, soggetto a piccole eccitazioni del tipo del rumore bianco, descriventi le fluttuazioni termiche, è stata eseguita da R. G. Williams, che ne ha riferito sul «SIAM Journal on Applied Mathematics» nel 1981. Lo stato stazionario è posto all'origine, e la figura mostra i risultati per cinque traiettorie di fuga. Sebbene il sistema trascorra la maggior parte del tempo nelle immediate vicinanze dello stato stazionario, effettua anche deviazioni erratiche, allontanandosene. Il moto di diffusione contrario al flusso non è un processo lento; copre distanze finite in intervalli di tempo relativamente corti. Se il sistema non raggiunge la frontiera circolare, viene attratto dalle vicinanze dello stato stazionario, dove inizierà un'altra escursione a un tempo successivo. Il risultato conferma la previsione teorica che le traiettorie si stringono attorno alla traiettoria a tempo rovesciato del sistema deterministico, che collega lo stato stazionario con lo stato della frontiera più vicino avente massima entropia, se diminuisce sufficientemente l'intensità del rumore.

Il problema che abbiamo descritto è quello di una particella che diffonde in direzione contraria al flusso. Questo problema diviene fisicamente più interessante se sono presenti più stati stazionari del moto. In tal caso esso può rappresentare la fuga di una particella da una buca di potenziale divenendo così un modello per le reazioni chimiche, per la diffusione nei cristalli, per le transizioni nelle giunzioni Josephson, così come per l'attraversamento delle barriere di potenziale da parte di particelle quantomeccaniche (effetto «tunnel»).

La presenza di fluttuazioni termiche casuali introduce un limite superiore alla precisione con la quale è possibile specificare un dato stato macroscopico di un sistema. Agli inizi degli anni cinquanta cominciarono intense ricerche per lo studio delle equazioni differenziali stocastiche del tipo di quella di Langevin; ne è risultato un nuovo tipo di calcolo, che prende il nome dal suo ideatore: Itô. Il calcolo stocastico di K. Itô è basato sull'osservazione che il cammino di una particella browniana è molto irregolare, se osservato con sufficiente dettaglio. (Abbiamo trascurato questo particolare nel formulare l'equazione di Langevin, che rimane comunque formale.) Ne consegue che, nel moto browniano non lo spostamento Δx, ma il suo quadrato è proporzionale a Δt, il coefficiente di proporzionalità essendo proprio due volte la costante di diffusione, D. In più è noto, fino dal 1933, che il moto browniano dà luogo alla relazione di indeterminazione:

$$\Delta x^2 \gtrsim 2D\Delta t$$

in diretta analogia con la relazione di indeterminazione vigente nella meccanica quantistica:

$$\Delta x^2 \gtrsim \frac{h}{m}\Delta t$$

in cui h è la costante di Planck e m è la massa della particella. Si può quindi osservare che in meccanica quantistica il rapporto $h/2m$ svolge lo stesso ruolo del coefficiente D nella teoria dei processi stocastici, e il calcolo stocastico può essere applicato in modo formale alla meccanica quantistica, sebbene non vi sia alcun processo di diffusione reale. La relazione di indeterminazione del moto browniano sta a significare che dobbiamo modificare le usuali regole di differenziazione (o di integrazione). Per ottenere il differenziale di una funzione è ora necessario prendere due termini nello sviluppo in serie di Taylor e sostituire il termine quadratico Δx^2 con il suo valor medio $2D\Delta t$.

La relazione di indeterminazione del moto browniano è una manifestazione delle correlazioni statistiche fra stati di non equilibrio attraverso i quali il sistema passa a istanti di tempo successivi.

La presenza di fluttuazioni termiche rende necessaria la transizione verso un'interpretazione probabilistica. La termodinamica statistica offre una connessione naturale fra la probabilità di uno stato e la sua entropia. Perciò l'entropia, anziché l'energia libera, ha un ruolo privilegiato quando vengono introdotti concetti probabilistici in termodinamica. Boltzmann fu il primo a riconoscere la connessione fra probabilità e entropia. Sfortunatamente visse in un'epoca in cui il determinismo della meccanica classica pervadeva le scienze naturali e soltanto dopo la sua morte Einstein riprese le sue idee e le applicò con tanto successo al moto browniano. Questo avvenimento, insieme alla spiegazione di Max Planck della radiazione del corpo nero, annunciò la nascita dell'era atomica.

Einstein mise in relazione la densità di probabilità $P(x)$, per una fluttuazione spontanea in uno stato di non equilibrio, con la diminuzione dell'entropia ΔS secondo la relazione:

$$P(x) \circ \exp [\, \Delta S(x) \,].$$

Questo implica che la probabilità, per differenti stati di non equilibrio, è proprio il prodotto delle probabilità individuali, confermando la proprietà di additività dell'entropia. Possiamo apprezzare la formula di Einstein per una fluttuazione spontanea dall'equilibrio come una forma limite, valida per tempi lunghi, qualora, cioè, le correlazioni statistiche abbiano avuto il tempo di attenuarsi, o al limite di piccole fluttuazione termiche. In generale vi sarà un'entropia addizionale generata dalle correlazioni statistiche fra stati di non equilibrio. Gli andamenti medi delle entropie di stato, opposti a quelli delle entropie correlazionali per stati di non equilibrio, sono schematizzati nella figura a pagina 29 in alto. Si può vedere che la presenza di correlazioni statistiche distrugge le proprietà termodinamiche di additività che caratterizzano l'equilibrio e che alla formula di Einstein per la densità di probabilità bisogna aggiungere una densità supplementare che tenga conto delle probabilità di transizione fra stati di non equilibrio.

Noi abbiamo derivato un'espressione generale per la densità di probabilità di transizione $P(x \rightarrow y)$, relativa a due stati di non equilibrio x e y; essa è data dall'analogo cinetico della formula di Einstein:

$$P(x \rightarrow y) \circ \exp [\, {}^1/_2 \,(\Sigma + \Delta S) \,]$$

in cui Σ è l'entropia congiunta e ΔS è la differenza di entropia fra i due stati di non equilibrio. L'entropia congiunta ha la proprietà di ridursi, dopo lungo tempo, alla somma delle entropie, e quest'ultima formula si riduce alla formula di Einstein.

Il principio generale della termodinamica del non equilibrio che governa l'evoluzione dei processi irreversibili verso l'equilibrio è:

$$\frac{d\overline{\Sigma}}{dt} \leqslant O$$

in cui la sopralineatura denota l'operazione di media. La diseguaglianza afferma che le correlazioni statistiche fra stati di non equilibrio sono, in media, una funzione decrescente del tempo. Usando le parole di Lars Onsager, essa esprime il fatto che tutti i sistemi fisici tendono a dimenticare il loro passato. Infine, la diseguaglianza può essere considerata come l'analogo stocastico del celebre teorema H di Boltzmann. Essa fornisce un meccanismo fisico per la conversione del nostro analogo cinetico nella formula di Einstein, al limite dei tempi lunghi.

Si può comprendere la differenza fra il modello deterministico (equazione diffenziale del primo ordine) e quello stocastico (equazione di Langevin) dei sistemi dinamici considerando che l'entropia è una funzione della variabile x che è lo scarto dall'equilibrio di una variabile di stato. Utilizzando l'equazione differenziale deterministica si scopre che l'entropia tenderebbe a zero per tempi lunghi, poiché per definizione x si annulla all'equilibrio (tempi lunghi). Se viene invece usata l'equazione di Langevin per calcolare l'entropia, troviamo che essa tende alla sua corretta forma d'equilibrio, al limite per tempi lunghi; questo dimostra che la presenza di fluttuazioni termiche casuali è fondamentale per lo stabilirsi della distribuzione d'equilibrio. In altre parole, l'equilibrio macroscopico corrisponde a uno stato medio attorno al quale il sistema fluttua a causa delle fluttuazioni termiche casuali. Infatti, mediando l'equazione di Langevin otteniamo l'equazione macroscopica fenomenologica della termodinamica dei processi irreversibili. Questa affermazione è stata enunciata per la prima volta da Onsager nella sua ipotesi della regressione delle fluttuazioni: la regressione delle fluttuazioni di non equilibrio obbedisce, in media, alle leggi fenomenologiche della termodinamica dei processi irreversibili. Il cammino che coincide con la soluzione deterministica della equazione cinetica (equazione differenziale del primo ordine) è chiamato «cammino termodinamico»; le fluttuazioni intorno al cammino termodinamico sono illustrate a pagina 29 in basso.

Il sistema progredisce verso un piccolo intorno dello stato di equilibrio caratterizzato dalla distribuzione statistica data dalla formula di Einstein; qui il sistema trascorre la maggior parte del suo tempo. Ora spostiamo il nostro asse del tempo a qualche istante lontano nel passato, affermando che il sistema invecchiato deve pur essere stato all'equilibrio molto tempo fa. A causa delle fluttuazioni termiche, il sistema avrà pure compiuto escursioni in stati a entropia inferiore di quella dello stato di equilibrio ma, con la diminuzione dell'intensità del rumore termico, il sistema evita gli stati nel dominio G con un'entropia inferiore di quello stato della frontiera B che possiede l'entropia massima, e attraverso il quale dovrebbe aver luogo l'uscita dal dominio G. In altre parole, con la diminuzione dell'intensità del rumore, il sistema non avrà energia sufficiente per visitare quella parte del dominio in cui l'entropia è inferiore a questo massimo sulla frontiera.

Nel caso di piccole fluttuazioni termiche, la prima uscita dalla frontiera avrà luogo, quasi certamente, in prossimità di quello stato U su B che rende massima l'entropia congiunta, Σ, soggetta alla condizione che il sistema sia stato nella vicinanza dello stato di equilibrio molto tempo fa. La fuga sarà avvenuta entro una piccola zona cilindrica, come un tubicino, attorno a un percorso che rende massima l'entropia congiunta. Sotto le stesse condizioni, questo cammino è l'immagine speculare nel tempo del percorso termodinamico, o deterministico, per la regressione delle fluttuazioni. Pertanto, il processo manifesta «una simmetria tra passato e futuro».

La linea piena mostra l'avvicinamento, entro un intorno Δ dello stato di equilibrio O, lungo il percorso più probabile o cammino termodinamico. Δ è una misura della grandezza delle fluttuazioni termiche attorno allo stato di equilibrio. Il cammino più probabile per l'uscita dalla regione G sarà entro un piccolo tubicino attorno al cammino che sia l'immagine speculare nel tempo del cammino più probabile per la regressione di una fluttuazione *(linea tratteggiata)*, al limite per l'intensità del rumore termico tendente a zero.

Questo comportamento è stato verificato con esperimenti di simulazione al calcolatore col metodo di Monte Carlo. Al diminuire dell'intensità del rumore termico, gli ultimi segmenti dei cammini che conducono al contorno si concentrano, quasi certamente, intorno a quella traiettoria che è l'immagine speculare del cammino più probabile per la regressione di una fluttuazione. Nella illustrazione a pagina 32 viene mostrato il processo chimico di Lotka, nel quale la rappresentazione delle curve integrali nel piano di fase indica che qualunque perturbazione del sistema «dinamico» causerà una reazione diretta alla restaurazione dell'equilibrio, che avrà luogo asintoticamente nel tempo, per piccole perturbazioni. A causa delle fluttuazioni termiche, è certo che il processo potrà sfuggire da qualunque dominio limitato contenente lo stato di equilibrio stabile. Per intensità molto piccole del rumore termico, la traiettoria di fuga convergerà con l'immagine speculare della traiettoria per il cammino più probabile della regressione di una fluttuazione.

Abbiamo chiamato l'immagine speculare del percorso termodinamico nel tempo col nome di «cammino antitermodinamico». Sebbene, a partire da un intorno dello stato di equilibrio, l'uscita dalla frontiera lungo il cammino antitermodinamico sia improbabile, anche l'uscita lungo qualsiasi altro percorso sarà improbabile, e in misura maggiore. Poiché, però, il tempo che il sistema trascorre nelle vicinanze dello stato di equilibrio è illimitato, l'uscita meno improbabile fino alla frontiera avrà luogo, prima o poi, e naturalmente nel modo più probabile. Ovviamente, il sistema avrebbe potuto raggiungere la frontiera lungo un'infinità

Lo schema della reazione chimica di Lotka viene usato anche come modello per fenomeni oscillatori che possono verificarsi negli ecosistemi. Il substrato iniziale A e il prodotto finale P sono mantenuti in quantità costanti dai flussi esterni. Gli intermedi X e Y variano nel tempo, in un modo descritto da equazioni deterministiche. Oscillazioni di queste concentrazioni nel tempo si verificano in situazioni lontane dall'equilibrio chimico, dove le reazioni inverse sono trascurabili in confronto alle reazioni dirette. In questo caso lo stato stazionario non è più stabile, nel senso che il sistema non agisce contro le perturbazioni che spostano il sistema dallo stato stazionario. Un fatto assai interessante è che anche quando ci troviamo nelle vicinanze dello stato di equilibrio, in cui si applica la legge dell'azione di massa, lo stato di equilibrio (situato all'origine del piano $[X-X_O, Y-Y_O]$) è instabile in presenza delle fluttuazioni termiche. Le equazioni cinetiche impiegate nella descrizione della reazione chimica sono ora del tipo dell'equazione di Langevin e, diversamente dal caso deterministico, il sistema compirà la sua uscita fuori da qualunque regione limitata contenente lo stato di equilibrio. Per piccole deviazioni dallo stato di equilibrio, la superficie entropica è convessa e ha il massimo di entropia S_O in corrispondenza dello stato di equilibrio. Le curve integrali, che sono ovunque tangenti al moto del sistema, sono mostrate in colore chiaro. Esse descrivono il moto del sistema nel suo avvicinarsi allo stato di equilibrio. In presenza di fluttuazioni termiche, il sistema farà la sua uscita da qualsiasi regione G, con contorno B, racchiudente lo stato di equilibrio. Nel limite di disturbi di intensità molto piccole, il cammino più probabile che il sistema seguirà nell'effettuare la sua uscita da G è la traiettoria del sistema deterministico, con il tempo rovesciato, che collega lo stato di equilibrio O con lo stato di massima entropia S^* sul contorno B, a patto che esista uno stato di massima entropia in B. Questa curva integrale col tempo invertito è rappresentata in colore più intenso e l'uscita si verificherà, quasi certamente, dallo stato U della frontiera che ha la massima entropia, al limite per il rumore termico tendente a zero. La traiettoria più probabile per l'uscita può essere trovata dalla condizione di stazionarietà dell'entropia congiunta, soggetta alla condizione che il sistema si sia trovato nello stato di equilibrio molto tempo prima.

di altri cammini, diversi da quello antitermodinamico. La probabilità del passaggio attraverso un tubicino, contenente uno di questi cammini qualsiasi, è infinitamente piccola se paragonata alla probabilità che il cammino avvenga all'interno di un tubicino contenente il percorso antitermodinamico. Occorre tuttavia notare che la somma delle probabilità di passaggio lungo gli infiniti tubicini possibili può anche essere più grande di quest'ultima.

Non è un caso che la meccanica quantistica ci dia una relazione di indeterminazione formalmente identica alla relazione di indeterminazione del moto browniano. Sin dalle prime formulazioni, si era intuito che la meccanica quantistica era incompleta: se fosse stato possibile specificare ulteriori variabili per descrivere qualche meccanismo interno fondamentale, ciò che ora è soltanto probabile sarebbe divenuto «certezza», rivelando così un determinismo soggiacente. Sebbene vi sia molta letteratura sulle teorie delle «variabili nascoste», nessuna di queste ha avuto successo nel dare un'interpretazione completamente soddisfacente di tutta la meccanica quantistica. Le formulazioni stocastiche della meccanica quantistica sono una via di mezzo fra la meccanica quantistica tradizionale e la teoria delle variabili nascoste, ipotizzando interazioni casuali fra le particelle quantistiche e il mezzo ipotetico nel quale esse si muovono. L'idea che una particella quantistica possa essere soggetta a un moto browniano classico è molto suggestiva, anche se prima sarà necessario riconciliare quest'idea col fatto che le traiettorie delle particelle quantistiche non sono osservabili, poiché qualsiasi tentativo di osservare la particella quantistica richiede un'interazione che perturba il sistema osservato. A questo punto ci piacerebbe credere che le probabilità siano reali, positive e normalizzabili, e che prevedano la frequenza con cui si verificano gli eventi reali. Per esempio, potremmo assegnare una distribuzione di probabilità a tutte le traiettorie possibili di una particella browniana. In meccanica quantistica, però, non vi è motivo per supporre l'esistenza di una distribuzione positiva e reale di probabilità per una data traiettoria, non potendosi eseguire più di una misura sullo stesso processo. (Una singola misura sulla particella causa la realizzazione di un evento osservabile, per il quale otteniamo una densità di probabilità reale che è sia positiva sia normalizzata.)

A differenza della meccanica classica, la meccanica quantistica è selettiva, nel senso che fa distinzione fra ciò che è fisicamente osservabile e ciò che è matematicamente «misurabile». Così la meccanica quantistica stocastica associa un processo, descritto da un'equazione del tipo di quella di Langevin, a ogni stato quantistico dinamico, e tutte le medie sono eseguite con una misura complessa della probabilità. La teoria è di per sé selettiva poiché risultano fisicamente osservabili solo quantità medie che siano reali. L'influenza dei campi esterni sul sistema entra attraverso la definizione del moto di deriva, e la formulazione usuale della meccanica quantistica mette in relazione il campo esterno con la funzione d'onda. In questo senso si può dire che la funzione d'onda determina lo stato del sistema, poiché specifica il moto tramite la deriva. Quindi non v'è alcuna differenza fisica fra il moto classico browniano e la meccanica quantistica; la sola differenza è di ordine matematico: l'uso di misure di probabilità complesse separa le osservazioni che hanno probabilità reali e positive di verificarsi da quelle non osservabili, che risultano poi essere quantità complesse. È un merito e un vantaggio dell'impostazione stocastica, che tutti i concetti probabilistici vengano introdotti in modo completamente classico. Se non altro il moto browniano ha fornito un modo per interpretare e capire da un punto di vista fisico cose note da tempo, ma astratte.

La freccia del tempo

Perché il tempo scorre in un'unica direzione? La risposta non va cercata nelle leggi di natura, che ben difficilmente riescono a distinguere tra passato e futuro, ma nelle condizioni prevalenti all'origine dell'universo

di David Layzer

Sembra facile distinguere il passato dal futuro, dato che la memoria ci fornisce una registrazione del passato, mentre del futuro non abbiamo una conoscenza certa. Tuttavia nell'interpretazione degli eventi secondo le leggi fondamentali della fisica la distinzione tra passato e futuro quasi scompare. La nostra percezione intuitiva del mondo ce lo rappresenta come qualcosa di esteso nello spazio ma « in sviluppo » nel tempo. Considerato in scala atomica il mondo è un continuo quadridimensionale esteso nello spazio e nel tempo. Intuitivamente diamo un significato particolare a un momento specifico, il presente, che vediamo come la cresta di un'onda che trasforma continuamente la potenza in atto e lascia nella sua scia il morto passato. La microfisica non privilegia nessun momento del tempo e fa solo una leggera differenza tra la direzione del passato e quella del futuro.

La nostra percezione intuitiva del mondo come qualcosa che si sviluppa nel tempo non può però essere abbandonata come qualcosa di meramente soggettivo, dato che ha una controparte oggettiva in una serie di processi biologici, geologici e astronomici. Aspetti evidenti sono i processi fisiologici che sono alla base della memoria, della crescita, dello sviluppo e della differenziazione degli organismi viventi, processi evolutivi, dove variazioni casuali associate alla selezione naturale hanno prodotto un numero enorme e in crescita costante di forme viventi organizzate in modo sempre più complesso. La crosta terrestre reca i segni di mutamenti evolutivi durati 4 miliardi e mezzo di anni, e le superfici piene di crateri della Luna e di Marte testimoniano il passaggio di un analogo periodo di tempo. Le stelle normali, le giganti rosse, le supernove e le nane bianche rappresentano stadi diversi nel ciclo evolutivo della vita di una singola stella. Infine la recessione delle galassie più lontane suggerisce l'idea che l'intero universo sia il prodotto di un processo evolutivo, che abbia avuto origine, in un tempo finitamente lontano, da uno stato indifferenziato estremamente denso.

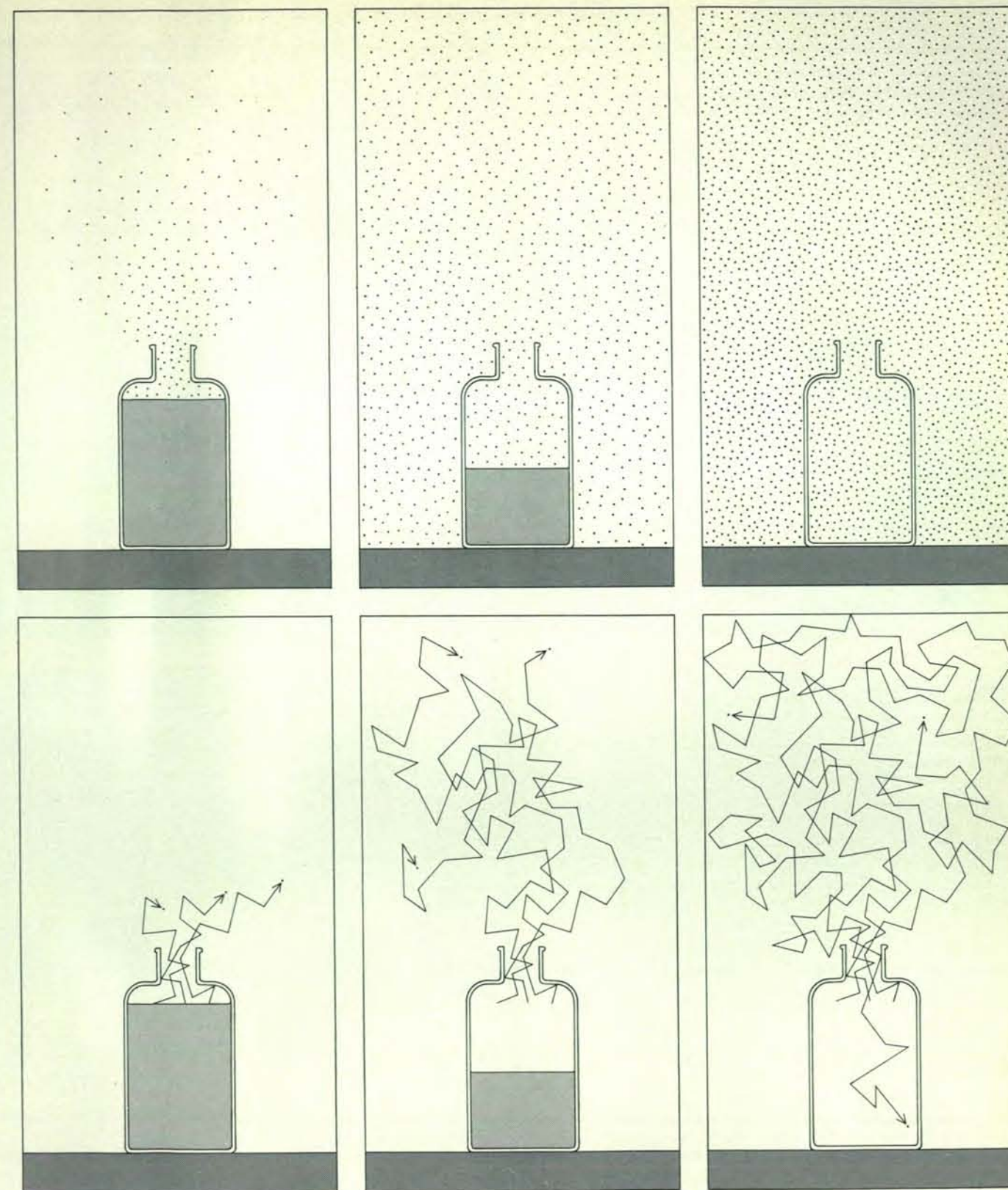
Tutti questi processi hanno qualcosa in comune: generano ordine, ossia informazione, trasformano uno stato più semplice in uno più complesso. Per usare un'espressione di Sir Arthur Eddington, indicano in che direzione è orientata la « freccia del tempo », definiscono quella che chiamo la freccia « storica » del tempo.

Paradossalmente è possibile definire la direzione del tempo per mezzo di una classe diametralmente opposta di processi, quelli che distruggono l'informazione e che generano disordine. Se lascio cadere una zolletta di zucchero in una tazza di tè caldo e mescolo, la concentrazione spaziale delle molecole di zucchero, il movimento del tè e la differenza di temperatura tra il tè e ciò che lo circonda costituiscono l'informazione macroscopica, ossia l'ordine. Col dissolversi dello zucchero, il raffreddarsi del tè e il cessare dei suoi movimenti l'informazione svanisce gradualmente. I processi irreversibili che distruggono l'informazione macroscopica (nel nostro esempio la diffusione molecolare, la viscosità e la conduzione del calore) sono manifestazioni della seconda legge della termodinamica. Tale legge stabilisce che tutti i processi naturali generano entropia, essendo l'entropia una misura del disordine. La distruzione irreversibile dell'ordine macroscopico definisce quella che ho chiamato freccia « termodinamica » del tempo.

A livello microscopico non si osserva né la freccia storica né quella termodinamica. Il movimento di una singola molecola di zucchero o di tè non genera né informazione né entropia. Il concetto di « ordine » è di tipo macroscopico, è una proprietà dei sistemi costituiti da molte particelle che non ha alcun senso quando sia applicata a singoli atomi o molecole. Nella fisica delle particelle elementari il mondo cambia senza evolversi.

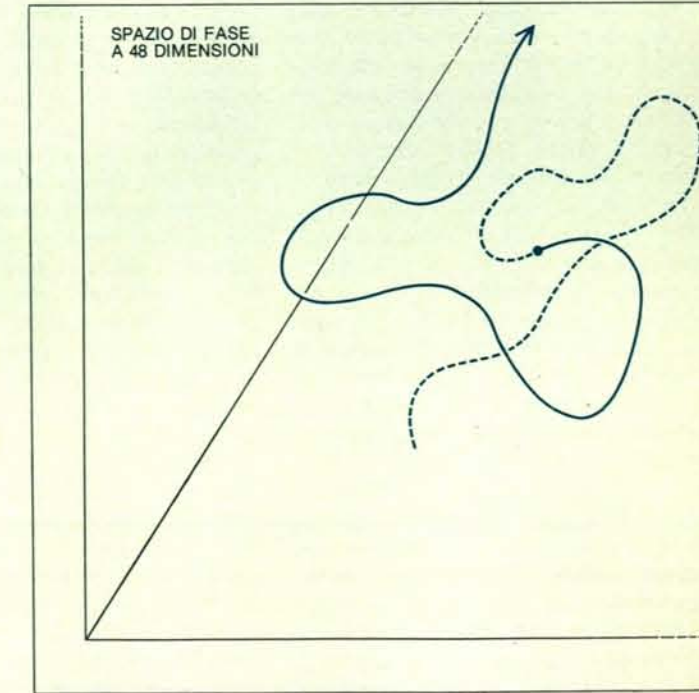
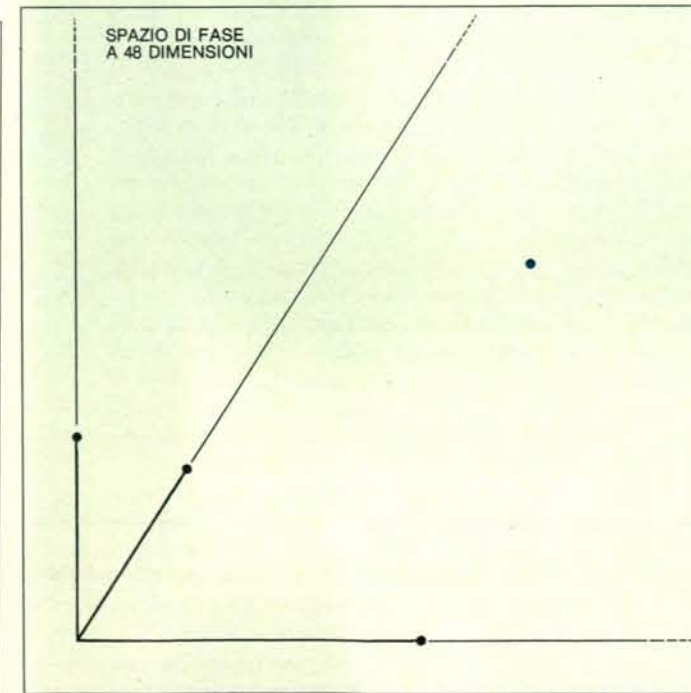
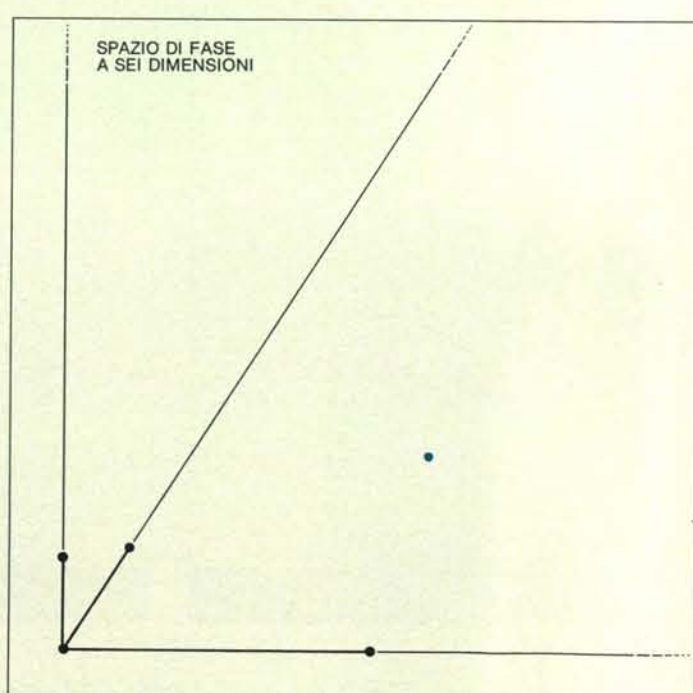
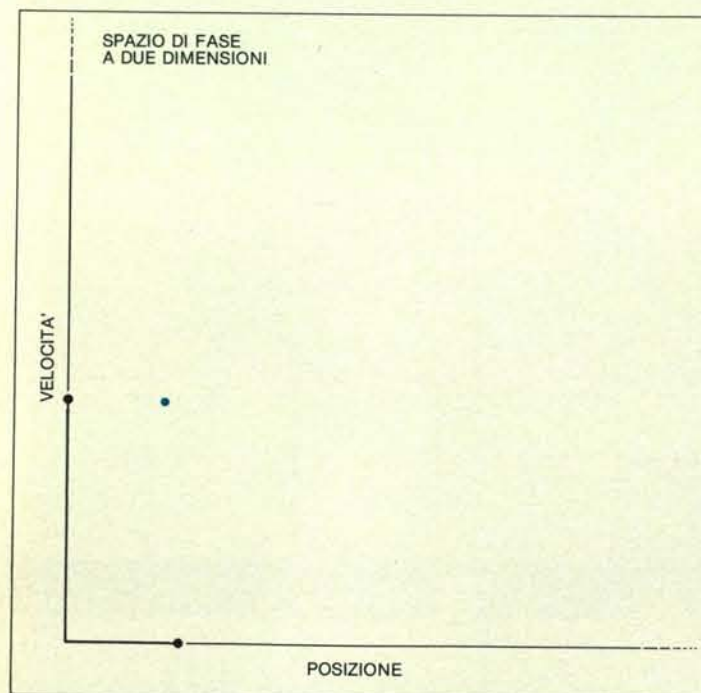
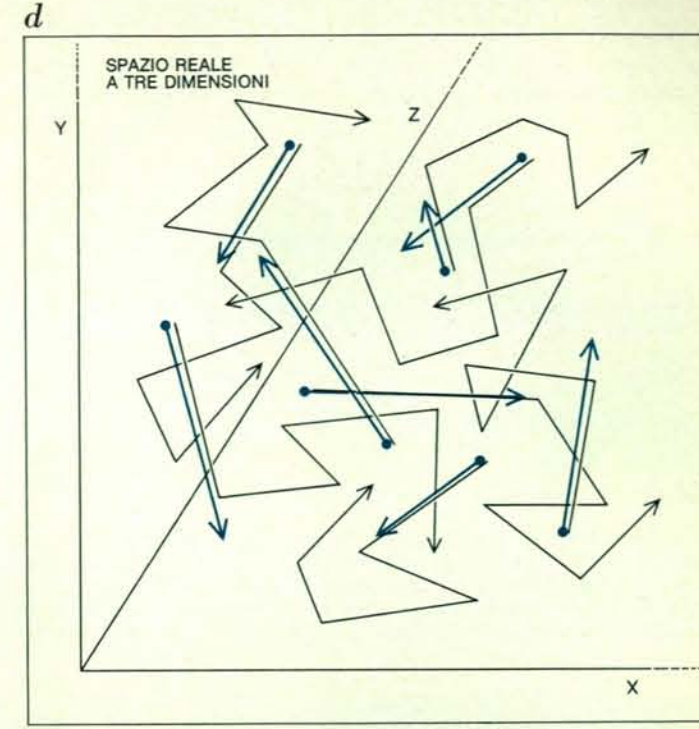
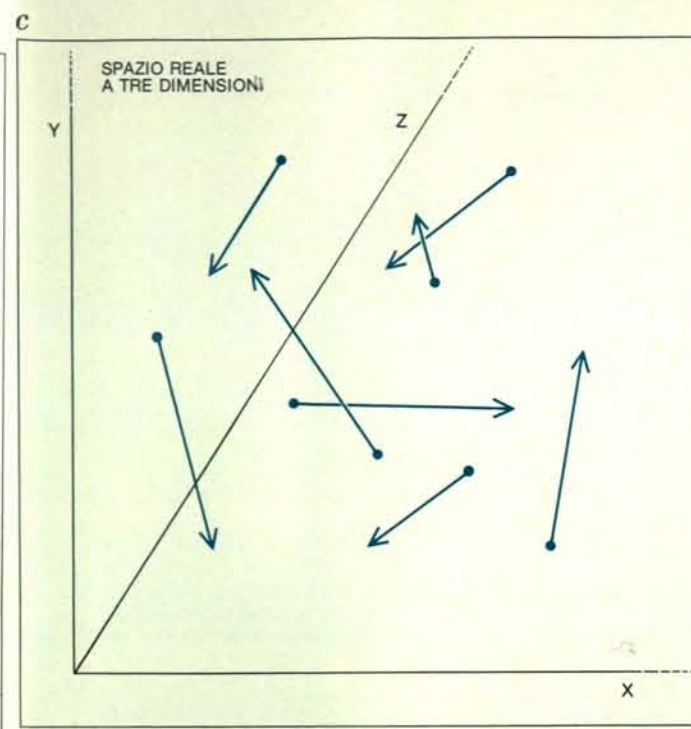
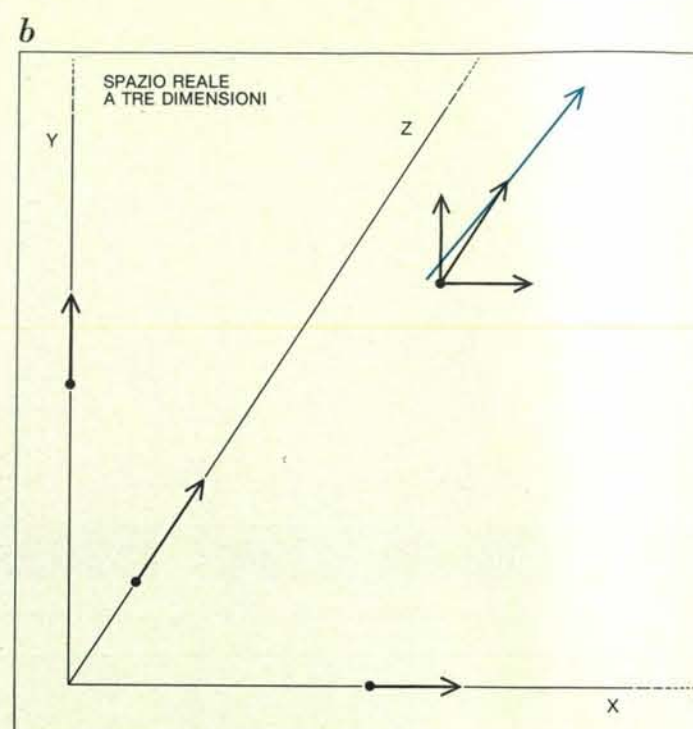
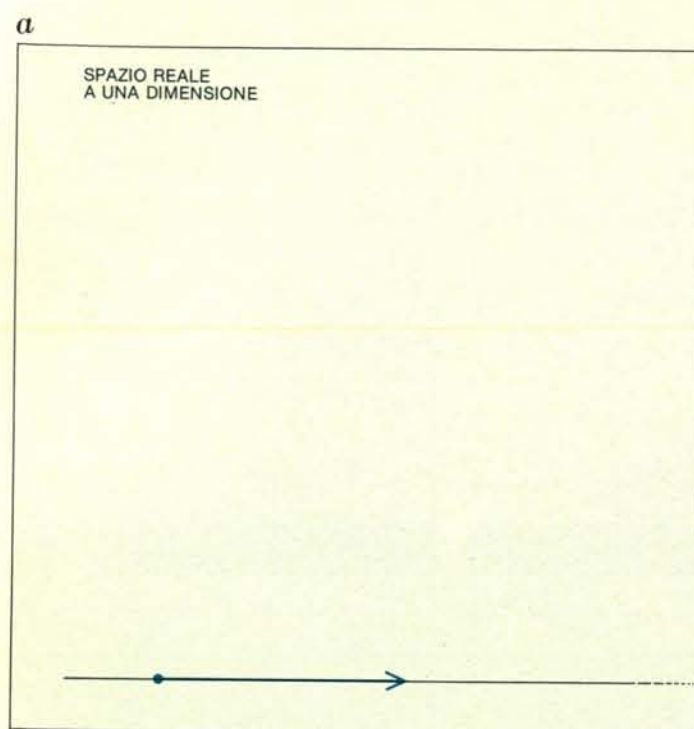
Io sostengo che né la concezione macroscopica del mondo come un sistema che degenera verso un disordine completo, né quella microscopica di un sistema di campi e particelle interagenti che muta senza evolversi, sono richieste dalle leggi fondamentali della fisica. La mia proposta consiste nel considerare entrambi i punti di vista come derivati da assunzioni ausiliarie circa la natura e l'origine dell'universo, e quindi nel sostituire tali assunzioni con altre, a mio parere più semplici, ugualmente in accordo con l'osservazione empirica. Il modello di universo che si ricava, sebbene differisca da quello accettato dalla maggior parte dei fisici, risolve l'apparente contraddizione tra freccia storica e termodinamica del tempo, riconciliandole entrambe con il carattere quasi simmetrico rispetto al tempo delle leggi fisiche a livello microscopico. Questa teoria implica che il mondo sia in espansione nel tempo e contemporaneamente spiega come diventi a ogni istante più complesso e ricco d'informazione.

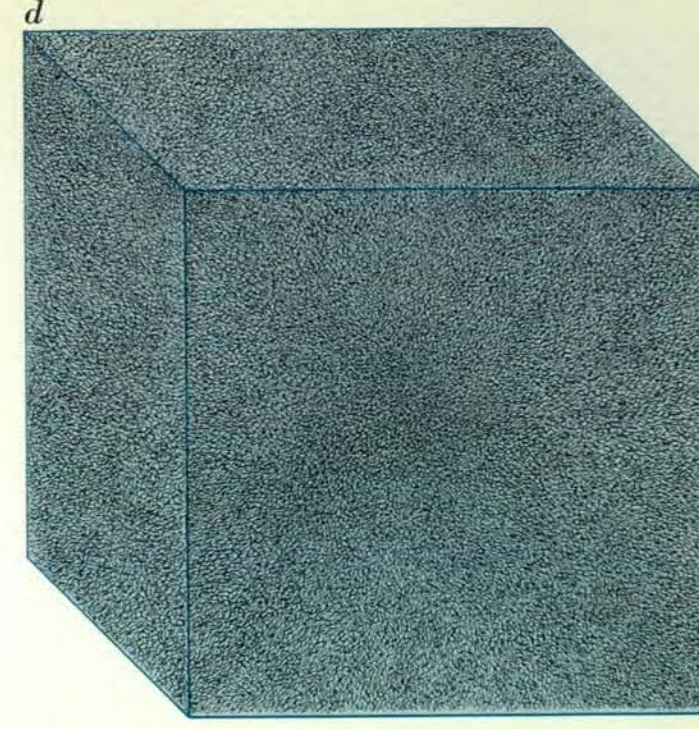
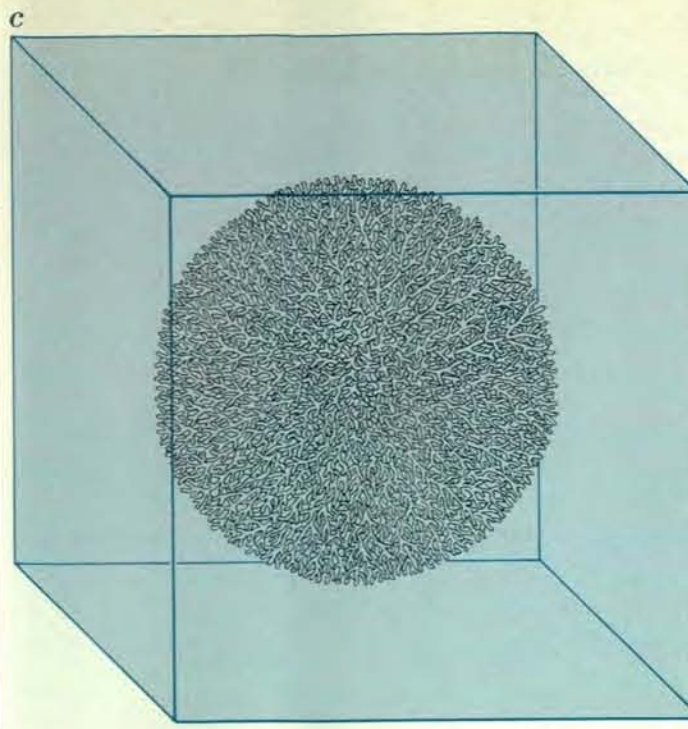
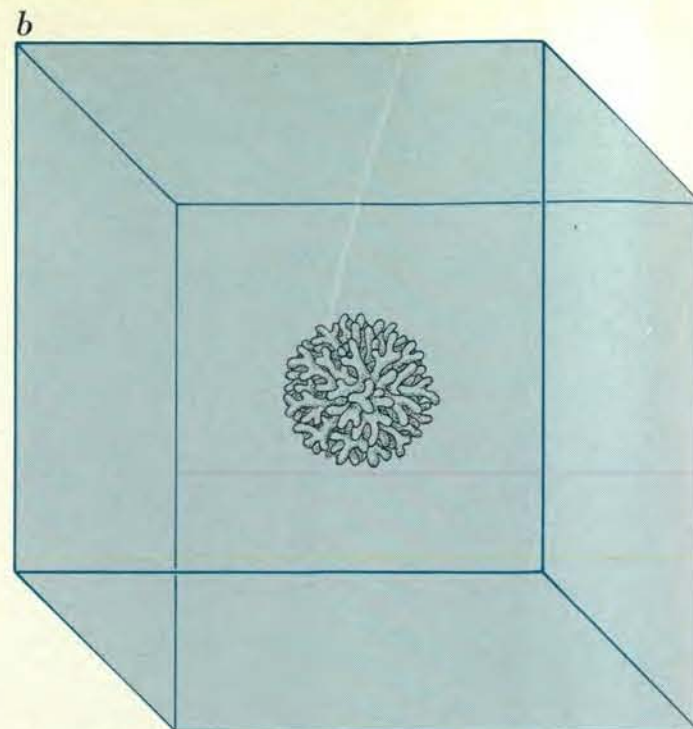
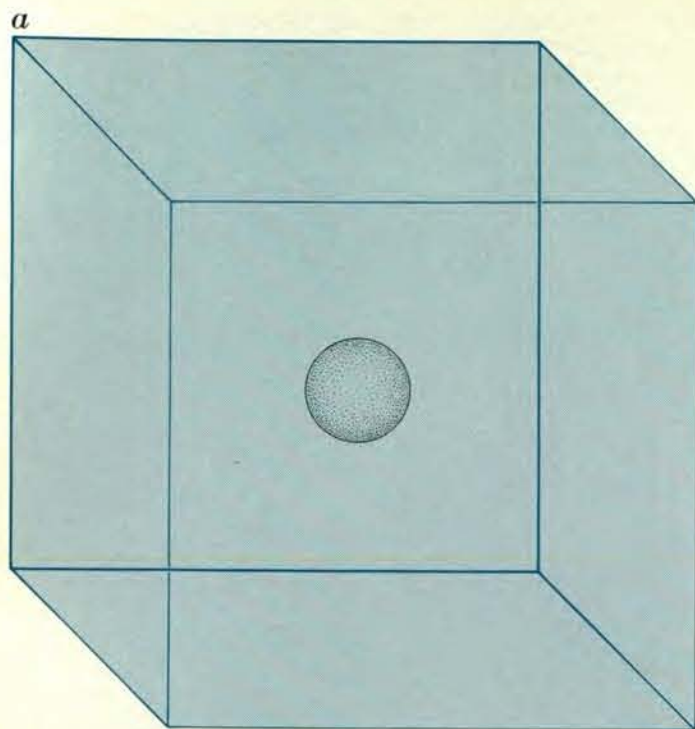
L'irreversibilità

La freccia storica e quella termodinamica derivano entrambe da processi che si svolgono sempre nella medesima direzione, il cui accadere è irreversibile. Che cosa rende irreversibili questi processi? Tutti i fenomeni possono essere descritti in ultima analisi come effetto dell'interazione di particelle elementari. Se le leggi che governano queste interazioni non distinguono tra passato e

Un esperimento mentale basato sulla diffusione del profumo permette di formulare conclusioni in apparenza paradossali: il processo nel suo complesso si sviluppa sempre nella medesima direzione, ma è costituito da eventi microscopici che presi singolarmente sono completamente reversibili. In una stanza ipotetica perfettamente isolata si apre una bottiglia di profumo. I disegni in alto, osservati da sinistra a destra, mostrano le molecole che si distaccano dalla superficie del liquido e che piano piano riempiono la stanza, fino all'evaporazione completa del liquido. Letti nel senso opposto i disegni rappresentano un processo mai osservato in natura: tutte le molecole entrano spontaneamente nella bottiglia e si condensano. Nei disegni in basso lo stesso processo è visto a livello microscopico. Le singole molecole lasciano la superficie del liquido e seguono complicate traiettorie a zig-zag diffondendosi nella stanza. Questa successione di eventi potrebbe benissimo svolgersi nell'altro senso, dato che, se ogni molecola invertisse la propria direzione, ripercorrendo la medesima strada, le molecole tornerebbero tutte nella bottiglia. In questo modo la molecola obbedirebbe a tutte le leggi della fisica e quindi sarebbe impossibile, esaminando il percorso di una singola molecola, capire se stiamo assistendo al processo normale oppure a quello che si svolge alla rovescia.

futuro, qual è il motivo dell'irreversibilità che osserviamo nel mondo macroscopico?

Una possibile spiegazione potrebbe trovarsi nel fatto che anche le leggi della microfisica non sono perfettamente simmetriche rispetto al tempo. L'esistenza di un'asimmetria temporale a livello di particelle subatomiche può essere confermata dal processo di decadimento di un mesone *K* neutro. Uno dei diversi possibili processi di decadimento di tale particella sembra violare alcune simmetrie della natura, e l'usuale interpretazione di questo fatto è che la simmetria violata sia quella dell'irreversibilità del tempo (si veda l'articolo *Esperimenti sull'inversione del tempo* di O. E. Overseth in « Le Scienze », n. 17, gennaio 1970). Tuttavia questa violazione è molto debole, dato che la si osserva in meno dell'uno per cento dei casi. Inoltre i mesoni *K* sono presenti solo in esperimenti di fisica delle alte energie; non sono ordinari costituenti della materia e non giocano alcun ruolo nei processi macroscopici che definiscono la freccia storica e quella termodinamica.

Se le radici dell'irreversibilità non si trovano nelle leggi che governano gli eventi microscopici, significa che andranno cercate nei vincoli che condizionano il verificarsi di tali eventi. Leggi e vincoli sono aspetti complementari della descrizione fisica della natura. Le leggi descrivono le regolarità sottostan-

Il concetto di spazio di fase è utilizzato per rappresentare lo stato dinamico di un sistema di particelle. Ogni particella è descritta da un vettore (*frecce colorate*) che definisce la posizione e la velocità. Per una particella in un universo unidimensionale (*a*) sono sufficienti due numeri per specificare questi parametri e lo stato della particella può essere rappresentato in uno spazio di fase a due dimensioni. Ogni possibile stato della particella corrisponde a qualche punto nello spazio di fase. Una particella libera di muoversi in uno spazio a tre dimensioni (*b*) richiede sei numeri per la specificazio-

ne del suo stato dato che sia la posizione sia la velocità hanno componenti su tre assi. Lo spazio di fase corrispondente deve perciò avere sei dimensioni. Dato che non è possibile costruire uno spazio reale con più di tre dimensioni, lo spazio di fase è qui rappresentato con una « fetta » tridimensionale dello spazio a sei dimensioni. In un sistema composto da molte particelle occorrono sei numeri per specificare lo stato di ogni particella, cosicché lo spazio di fase corrispondente deve avere un numero di dimensioni uguale a quello delle particelle moltiplicato per sei. Per esempio, un sistema di otto particelle (*c*) potrebbe essere rappresentato da un punto in uno spazio di fase a 48 dimensioni. L'informazione necessaria per specificare lo stato è condensata tutta nella collocazione di questo singolo punto e ogni stato possibile corrisponde a un unico punto nello spazio di fase a 48 dimensioni. Gli assi della fetta tridimensionale sono scelti in modo arbitrario tra i 48. Con l'evolversi del sistema di particelle (*d*) il suo stato dinamico cambia e questo mutamento si riflette nel movimento del punto che lo rappresenta nello spazio di fase. Il percorso del punto, sia nel passato sia nel futuro, è completamente determinato dalla sua posizione iniziale e quindi la storia dinamica del sistema può essere completamente predetta nei dettagli. Inoltre il punto nello spazio di fase può seguire lo stesso percorso in entrambe le direzioni e ciò rappresenta la completa reversibilità dei movimenti delle particelle nello spazio reale. Anche in questo caso nella nostra rappresentazione abbiamo scelto tre dimensioni qualsiasi tra le 48 a nostra disposizione.

La rappresentazione probabilistica di un sistema composto da molte particelle raffigura in modo più realistico il comportamento del sistema. La probabilità è mostrata come un fluido nello spazio di fase; la massa del fluido in una regione rappresenta la probabilità che il punto che rappresenta lo stato del sistema si trovi in quella regione. Nell'esperimento sulla diffusione del profumo tutto il fluido probabilistico è concentrato all'inizio in un piccolo volume, dato che tutte le molecole sono racchiuse nella bottiglia. La forma del fluido è di fatto un'ipersfera a $6n$ dimensioni, dove n è il numero delle molecole di profumo; nel disegno è rappresentata da una sfera a tre dimensioni (*a*). Con l'evolversi del sistema il fluido deve migrare in regioni dello spazio di fase più lontane, ma dato che le traiettorie delle particelle sono determinate il fluido risulta incomprimibile. La sua espansione non è come quella di un gas, ma assomiglia all'estroflessione di « tentacoli » (*b*), sempre più sottili e numerosi (*c*) col crescere degli stati possibili del sistema. Alla fine tutto l'ipervolume è pieno di sottili rami del fluido, il cui volume totale rimane costante. Da un punto di vista macroscopico la distribuzione del fluido sembra ora uniforme, ma non è così a un esame ravvicinato (*d*).

ti i fenomeni, sono poco numerose e si applicano a un dominio molto vasto. I vincoli servono per selezionare, nell'ambito degli eventi governati da una legge, il particolare fenomeno che interessa. Le leggi definiscono l'ambito del possibile, i vincoli ciò che è reale o rilevante. I vincoli possono prendere la forma di condizioni iniziali, condizioni limitanti o condizioni di simmetria.

Per illustrare come leggi e vincoli contribuiscano insieme a dar forma ai fenomeni, consideriamo i movimenti dei pianeti del sistema solare. In base alla legge di gravitazione di Newton si potrebbero calcolare tutte le posizioni passate e future dei pianeti conoscendo le loro posizioni e velocità in un dato momento. La legge di Newton spiega perché ogni pianeta segua un'orbita ellittica con il Sole in uno dei fuochi, perché il movimento di una linea che connetta il Sole con un pianeta generi aree uguali in tempi uguali e perché i quadrati dei periodi orbitali dei pianeti siano proporzionali ai cubi dei diametri delle loro orbite. La legge in questione potrebbe spiegare questi fatti per ogni sistema planetario. D'altra parte la legge di gravitazione non spiega perché le orbite dei pianeti sono quasi circolari, perché i piani orbitali sono quasi coincidenti o perché tutti i pianeti girano intorno al Sole nella stessa direzione. Secondo quanto riconobbe lo stesso Newton, tali regolarità devono avere origine dalle condizioni iniziali.

Per spiegare queste regolarità occorrerebbe una teoria della formazione dei pianeti. Tale teoria non potrebbe fornire in dettaglio le condizioni iniziali del sistema solare, ma potrebbe specificare certe proprietà statistiche dei sistemi primordiali da cui hanno avuto origine i sistemi planetari come il nostro. Anche questa teoria dovrebbe muovere però da certe condizioni iniziali particolari, che a loro volta potrebbero mostrare certe regolarità statistiche che spingerebbero la spiegazione teorica a un livello più profondo. In questo modo saremmo condotti a formulare una serie di problemi cosmogonici sempre più generali le cui soluzioni darebbero origine a spiegazioni sempre più generali delle regolarità statistiche dell'universo astronomico. Questa ipotetica catena di teorie cosmogoniche deve terminare alla fine nella specificazione di un insieme di vincoli riguardanti l'universo nel suo complesso, comprese certe condizioni iniziali. È proprio in questo insieme di condizioni vincolanti che possiamo trovare la radice dell'irreversibilità.

Informazione ed entropia

I processi che definiscono la freccia storica e la freccia termodinamica del tempo sono quelli che rispettivamente generano informazione e entropia. Come dimostrò nel 1946 Claude E. Shannon del Massachusetts Institute of Technology, l'informazione è una proprietà delle descrizioni statistiche dei sistemi fisici. Essa viene misurata in bit, o cifre binarie, dove per bit si intende la quantità d'informazione necessaria per decidere tra due possibilità egualmente probabili. L'informazione può anche essere considerata come una proprietà dei sistemi fisici stessi, una misura del loro livello di organizzazione. Un teorema fondamentale dovuto a Shannon dimostra che il contenuto d'informazione di un sistema è pari al minimo numero di bit necessario per codificare una descrizione statistica completa del sistema.

Il concetto di entropia è strettamente connesso a quello di informazione. La entropia è stata definita per la prima volta (da Rudolf Clausius e da Lord Kelvin) nel contesto della termodinamica e consiste nella misura dello spostamento di un sistema dall'equilibrio termodinamico; nello stato di equilibrio l'entropia assume il valore massimo per dati valori di temperatura e di densità.

Utilizzando una formula derivata per la prima volta da Ludwig Boltzmann e da J. Willard Gibbs, Shannon definì un concetto di entropia all'interno della teoria dell'informazione, consistente nella misura del grado di incertezza associato alla descrizione statistica di un sistema. L'entropia termodinamica di Kelvin e Clausius e quella statistica di Boltzmann, Gibbs e Shannon godono di identiche proprietà matematiche: sono aspetti differenti di un medesimo concetto.

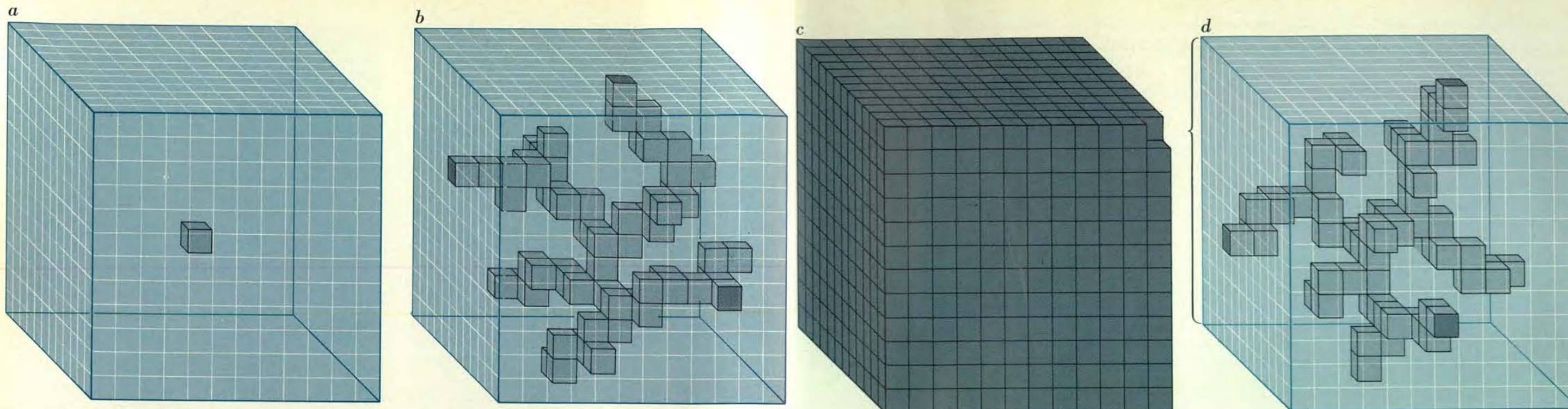
Entropia e informazione sono connesse da una semplice legge di conservazione, che stabilisce che la somma dell'informazione e dell'entropia è costante e uguale al massimo di informazione che si può avere dal sistema, ossia all'entropia sotto certe condizioni date. Esprimendo matematicamente questa legge, abbiamo che $H+I$=costante $=H_{max}=I_{max}$, dove H (la lettera greca eta) e I rappresentano i valori effettivi di entropia e di informazione, mentre H_{max} e I_{max} rappresentano i massimi valori possibili. Dunque un aumento d'informazione è sempre compensato da una uguale perdita di entropia.

Supponiamo che un sistema fisico disponga di otto (ossia 2^3) possibili stati. In notazione binaria potrebbero essere indicati dai numeri 000, 001, 010, 011, 100, 101, 110 e 111. La specificazione di uno stato particolare, per esempio quello indicato da 101, richiede tre cifre binarie, vale a dire la quantità d'informazione associata alla descrizione seguente: « Il sistema è chiaramente nello stato 101 ». Il grado di incertezza, ossia l'entropia, associato a questa descrizione è evidentemente nullo. All'estremo opposto, se non possedessimo alcuna informazione circa lo stato del sistema, saremmo costretti ad assegnare uguali probabilità a ciascuno degli otto possibili stati. In questo caso l'informazione è zero. Dato che la somma di entropia e di informazione è costante per il sistema, l'entropia deve ora essere di tre bit. In generale, se un sistema ha 2^r stati possibili, dove r è un numero intero, allora la massima quantità di informazione o entropia è uguale al logaritmo in base 2 di 2^r, ossia r.

Un esperimento mentale

Per i sistemi reali il numero dei possibili stati è molto grande ma non infinito. Il numero degli stati, e quindi anche la quantità massima d'informazione, sono limitati dal principio d'indeterminazione formulato da Werner Heisenberg. Secondo tale principio la conoscenza della posizione e del momento di una particella reale è limitata da una certa quantità irriducibile di indeterminazione, in modo che lo stato della particella non può essere specificato con maggior precisione di quella consentita dal principio d'indeterminazione. Ne deriva che qualsiasi sistema fisico può essere descritto con una quantità finita di informazione.

Per comprendere i rapporti tra cosmologia, entropia e informazione possiamo compiere un semplice esperimento mentale. In un angolo di una stanza in cui l'aria sia perfettamente calma viene aperta una bottiglia di profumo. Dopo qualche tempo un osservatore situato nell'angolo opposto della stanza riferisce di cominciare a sentire l'odore del liquido. Evidentemente le molecole di profumo si sono allontanate dalla superficie del liquido e dopo un certo numero di collisioni con altre molecole, seguendo un complicato percorso a zig-zag, hanno attraversato la stanza. Dopo un tempo sufficientemente lungo tutto il profumo sarà evaporato e le molecole di profumo saranno distribuite in modo uniforme nella stanza (*si veda la figura a pagina 27*).

L'esperienza e la seconda legge della termodinamica ci dicono che il processo è irreversibile: per quanto a lungo si aspetti le molecole di profumo non se ne torneranno mai spontaneamente nella bottiglia. Tuttavia non si tratta di un evento impossibile in linea di principio. Supponiamo di aver registrato l'intero esperimento su un film che ne riprendesse i dettagli microscopici, in modo da poter seguire individualmente ogni molecola nei suoi movimenti. Potremmo vedere allora come una singola molecola si distacca dalla superficie del liquido in seguito all'impulso ricevuto da una collisione. Il suo percorso av-

La teoria dell'informazione fornisce un'interpretazione quantitativa della distribuzione del fluido probabilistico. Lo spazio di fase è diviso in piccole celle di uguale ipervolume e nello stato iniziale si suppone che tutto il fluido sia contenuto in una cella (a). L'informazione richiesta per specificare questa distribuzione è pari al logaritmo in base 2 del numero delle celle. Con l'evolversi del sistema di particelle, il fluido passa a occupare sempre nuove celle (b) e alla fine la distribuzione è uniforme; ogni cella contiene un uguale volume di fluido (c). Lo stato del sistema è allora indeterminato e non occorre alcuna informazione per precisarlo. Ma se si esamina lo spazio di fase su scala più ridotta si trova che la distribuzione del fluido non è affatto uniforme. Dividendo ogni cella in altre celle (d) si dimostra che l'informazione necessaria per definire lo stato del sistema è restata invariata: è ancora il logaritmo in base 2 del numero delle celle. L'informazione dello stato iniziale sullo stato macroscopico del sistema è diventata informazione sullo stato microscopico. Si dimostra che se l'informazione microscopica è assente nello stato iniziale del sistema, allora tutta l'informazione macroscopica del sistema viene convertita in informazione microscopica.

viene tra collisioni continue con migliaia di altre particelle dell'aria, con le pareti del contenitore, con i muri della stanza, mutando ogni volta direzione e velocità. Dopo un certo tempo la ritroviamo ancora in movimento dall'altra parte della stanza. Se si proiettasse alla rovescia un film di questo tipo, si vedrebbero le molecole di profumo percorrere alla rovescia le loro complicate traiettorie per radunarsi tutte nella bottiglia e riunirsi a formare un liquido. Se isolassimo una singola molecola troveremmo che il suo percorso obbedisce a tutte le leggi della fisica, dato che le leggi che governano il moto delle molecole sono simmetriche rispetto al tempo. Nessuna caratteristica della traiettoria di una singola molecola è in grado di farci distinguere il film « vero » da quello alla rovescia. Perché allora respingiamo il film alla rovescia come registrazione di un fatto reale?

L'ovvia risposta è che nel film alla rovescia le condizioni iniziali sono molto particolari. Infatti all'inizio di questo film ognuna di una gran massa di molecole è su una traiettoria che alla fine la condurrà in un certo volume di spazio molto ristretto (la bottiglia di profumo) escludendo tutti gli altri volumi simili a sua disposizione (il resto della stanza). Uno stato iniziale di questo tipo è estremamente improbabile e ciò è stato considerato spesso una spiegazione sufficiente dell'irreversibilità dei processi termodinamici. Tuttavia è possibile spingere l'analisi più a fondo e chiedersi che cosa renda tanto improbabili queste condizioni iniziali.

Lo spazio di fase

Per continuare l'analisi abbiamo bisogno di strumenti adatti alla rappresentazione del mutamento dello stato dinamico di un sistema contenente un gran numero di particelle. Lo spazio di fase fa al caso nostro.

Lo stato dinamico di una singola particella è completamente descritto dalla sua posizione e dalla sua velocità. Per esprimere queste quantità abbiamo bisogno di sei numeri: tre coordinate per la posizione e tre componenti per la velocità. In coordinate cartesiane i numeri corrispondono a posizione e velocità lungo gli assi x, y e z. Questi sei numeri possono considerarsi come le sei coordinate di posizione di un punto in uno spazio a sei dimensioni, lo spazio di fase della particella. A ogni punto dello spazio di fase corrisponde un unico stato dinamico della particella nello spazio reale e il movimento della particella nello spazio reale si traduce in una curva nel suo spazio di fase tracciata dal suo punto di rappresentazione. Se conoscessimo la posizione e la velocità della particella in un istante potremmo predire ogni suo movimento successivo con la massima precisione; in altre parole, la storia dinamica della particella è completamente determinata dalle sue condizioni iniziali. Analogamente, nello spazio di fase l'intera curva è completamente determinata dal suo punto di origine. Inoltre il percorso del punto nello spazio di fase non può tornare su se stesso né biforcarsi (sebbene possa descrivere una curva chiusa). Infatti se tale percorso si intersecasse vi sarebbe uno stato della particella (quello rappresentato dal punto d'intersezione) con più di uno stato successivo e la storia dinamica della particella non sarebbe determinata in modo univoco.

Con la medesima tecnica è possibile descrivere un sistema chiuso composto da molte particelle interagenti. Lo stato dinamico di un sistema di n particelle è specificato da $6n$ numeri: le tre coordinate di posizione e le tre componenti della velocità di ognuna delle n particelle. Questi numeri possono essere considerati come le coordinate di un punto in uno spazio a $6n$ dimensioni; per descrivere il sistema dobbiamo specificare la collocazione di un singolo punto in questo spazio. Analogamente la storia dinamica del sistema è rappresentata da una curva nello spazio di fase che è determinata in modo completo e univoco dal suo punto di origine. Tale curva può avere una forma complicata o irregolare in relazione alle collisioni o a altre interazioni che si verificano tra le particelle, ma non può biforcarsi né tornare su se stessa.

La diffusione delle molecole di profumo nel nostro esperimento mentale è rappresentata da un'unica traiettoria in uno spazio di fase a $6n$ dimensioni, dove n è il numero delle molecole di profumo. (Si tratta di un numero molto grande: per un grammo di profumo n è pari circa a 6×10^{20}.) La traiettoria congiunge i punti che rappresentano lo stato iniziale e quello finale dell'esperimento, ma se potessimo esaminare questi punti non potremmo trovare alcuna differenza qualitativa tra di essi. Ognuno di essi può essere conseguenza dell'altro e la descrizione dei movimenti delle molecole è completamente reversibile.

L'analisi del nostro esperimento mentale compiuta per mezzo dello spazio di fase sembra abolire la freccia del tempo, imponendo un determinismo che non lascia spazio per la novità. Tuttavia si tratta di una descrizione dotata di una precisione che non è realistica. Si supponga infatti di possedere riguardo al sistema delle molecole di profumo e dell'aria una quantità d'informazione superiore a quella che di fatto è possibile avere. Di fatto non conosciamo le posizioni iniziali precise e le velocità delle 6×10^{20} molecole di profumo, anche tenendo conto delle limitazioni imposte dal principio d'indeterminazione. Tutto ciò che possiamo dire è che sono confinate all'inizio in un certo volume, la bottiglia, ma non possiamo certo specificare le coordinate del punto che rappresenta il sistema nello spazio di fase. Tutto quello che possiamo dire è che tale punto deve essere compreso all'interno di un piccolo volume, o « ipervolume », dello spazio di fase a $6n$ dimensioni.

Allo scopo di rappresentare questa informazione rimpiazziamo il punto nello spazio di fase con una bolla costituita da un fluido immaginario che riempia in modo uniforme il piccolo ipervolume corrispondente alla nostra effettiva conoscenza dello stato iniziale. Il fluido immaginario rappresenta la probabilità e la massa di fluido in una regione dello spazio rappresenta la probabilità che lo stato dinamico del sistema corrisponda a un punto situato all'interno di quella regione (*si veda la figura alle pagine 30 e 31*).

Come si diffonde il fluido probabilistico nello spazio di fase man mano che le molecole di profumo si diffondono nello spazio fisico? Si potrebbe pensare che si espanda in tutte le direzioni, come il profumo, e che riempia alla fine tutto l'ipervolume in modo più o meno uniforme. Di fatto il comportamento del fluido è alquanto diverso.

Dato che i movimenti delle molecole di profumo sono determinati in modo completo dal loro stato iniziale (anche se questo ci è ignoto), il fluido probabilistico deve restare una singola bolla continua. Se dovesse spezzarsi in due o più bolle separate ciò significherebbe che esiste una storia dinamica rappresentata da una traiettoria che si biforca, fatto che come abbiamo visto è impossibile. Inoltre il volume della bolla non può cambiare, dato che è proporzionale al numero di stati distinguibili permessi dal principio d'indeterminazione e tale numero non può cambiare finché ogni stato definisce un'unica storia dinamica.

Da queste considerazioni si può concludere che il fluido probabilistico è continuo e incomprimibile e che quindi il suo comportamento è più simile a quello di un liquido che a quello di un gas. La sua espansione nell'ipervolume non avviene per una diminuzione di densità, come nel caso di un gas, ma estroflettendo dei « tentacoli » che diventano sempre più lunghi e stretti e sempre più numerosi man mano che il sistema si evolve. Il processo è stato paragonato da Gibbs a quello con cui l'inchiostro di China si diffonde in acqua ferma.

Col diffondersi del fluido probabilistico tramite l'emissione di tentacoli sem-

pre più sottili, l'ipervolume totale occupato dal fluido resta costante ma la forma della regione occupata diventa sempre più complessa. Dopo un periodo sufficiente di tempo il fluido sembrerà distribuito in modo uniforme nell'intero ipervolume, tuttavia quando lo si esamini nei minimi particolari si troverà che la distribuzione è tutt'altro che uniforme. In questa descrizione del nostro esperimento mentale abbiamo rilevato una differenza enorme tra lo stato iniziale e quello finale. All'inizio il fluido probabilistico è confinato in una piccola regione dello spazio di fase che occupa in modo uniforme, mentre il resto dell'ipervolume è vuoto. Nello stato finale il fluido occupa l'intero ipervolume. Da un punto di vista macroscopico esso sembra distribuito in modo uniforme, ma in scala microscopica rivela una distribuzione assai poco uniforme.

Il flusso dell'informazione

La distinzione tra una distribuzione uniforme e una non uniforme del fluido probabilistico rappresenta una differenza qualitativa nel contenuto d'informazione del sistema. Per misurare l'informazione dobbiamo dividere la regione accessibile dello spazio di fase in piccole celle di eguale ipervolume. Per comodità supponiamo di ottenere 2^r celle, dove r è un numero intero (*si veda la figura alle pagine 32 e 33*). All'inizio il fluido probabilistico è confinato in una delle celle. L'informazione richiesta per specificare lo stato è semplicemente il numero di cifre binarie necessario per specificare una particolare cella. Il numero di bit richiesto è il logaritmo in base 2 del numero delle celle, ossia $\log_2 2^r = r$. In conclusione lo stato iniziale dell'esperimento mentale può essere rappresentato da r bit d'informazione.

Nello stato finale, quando il fluido è distribuito in modo uniforme, ognuna delle 2^r celle contiene lo stesso volume di fluido probabilistico. A quel livello di descrizione lo stato finale è completamente indeterminato e non occorre alcuna informazione per specificarlo. Nell'evoluzione del sistema sembra che sia scomparsa tutta l'informazione contenuta nello stato iniziale.

Se esaminiamo la distribuzione del fluido su una scala più ridotta riusciamo tuttavia a scoprire dove è andata a finire l'informazione. Se ogni cella contiene un volume uguale di fluido e se il volume totale del fluido non è mutato, allora all'interno di ogni cella il fluido probabilistico deve occupare solo $1/2^r$ del volume. Sebbene la densità del fluido non abbia subito variazioni, la forma della regione di spazio occupata è diventata ora molto complessa. Se si divide la cella in « microcelle » sufficientemente piccole si può dimostrare che l'informazione necessaria per specificare la distribuzione del fluido nell'intera regione dello spazio di fase è sempre $\log_2 2^r = r$. Quindi l'informazione macroscopica presente nello stato iniziale non è scomparsa ma si è semplicemente convertita in informazione microscopica nello stato finale.

Questa conclusione può essere generalizzata e precisata in modo esauriente. Comunque si decida di dividere lo spazio di fase in « macrocelle », possiamo definire l'informazione macroscopica come l'informazione necessaria per specificare l'insieme di probabilità associato a queste macrocelle; l'informazione necessaria per specificare la distribuzione del fluido all'interno delle macrocelle viene definita microscopica. Con l'evolversi del sistema chiuso di molecole, la quantità totale d'informazione necessaria per specificare la distribuzione del fluido probabilistico nello spazio di fase del sistema si mantiene a un livello costante, ma l'informazione macroscopica può essere convertita in informazione microscopica e viceversa.

Che cosa rappresentano questi due tipi di informazione? Possiamo identificare l'informazione macroscopica con la nostra conoscenza delle proprietà statistiche del sistema, e l'informazione microscopica con una conoscenza dettagliata dello stato delle singole molecole. In particolare l'informazione microscopica rappresenta la nostra conoscenza delle correlazioni esistenti tra le velocità delle particelle. Nel nostro esperimento mentale l'informazione microscopica era inizialmente assente, dato che nello stato iniziale non vi erano correlazioni tra le velocità delle molecole; la conoscenza della velocità di una molecola non ci avrebbe messo in grado di predire la velocità di qualsiasi altra molecola. Con l'evolversi del sistema le collisioni creavano delle correlazioni tra le velocità delle particelle e tutta l'informazione macroscopica presente si convertiva alla fine nell'informazione microscopica rappresentata da queste correlazioni.

Per certi tipi di sistemi fisici e sotto certe condizioni iniziali si può dimostrare che questo processo è inevitabile. Se l'informazione microscopica è assente all'inizio in un sistema composto da molte particelle interagenti, allora l'informazione necessaria per specificare lo stato macroscopico del sistema deve diminuire rapidamente finché si converte tutta in informazione microscopica. Teoremi di questo tipo sono stati dimostrati a partire dal 1946 (per particolari classi di sistemi fisici e per particolari definizioni del concetto di informazione microscopica) da Nikolai Bogolyubov, Leon C. P. van Hove, Ilya Prigogine, Radu Balescu, Mark Kac e altri.

Dato che l'informazione macroscopica invariabilmente diminuisce in quelle circostanze in cui l'entropia termodinamica aumenta, si è tentati di definire l'entropia termodinamica come infor-

a

b

Perturbazioni casuali provenienti dall'esterno del sistema di particelle tendono a dissipare l'informazione microscopica. In un sistema che non può comunicare con l'esterno (*a*) i percorsi delle particelle sono determinati e il fluido probabilistico nello spazio di fase in cui è rappresentato il sistema risulta incomprimibile. Tuttavia nessun sistema reale è veramente isolato. Le pareti di ogni contenitore trasmettono il calore e vi è inoltre l'interazione gravitazionale delle particelle con materia distante. Ne risulta che perturbazioni casuali distruggono tutta l'informazione sullo stato microscopico del sistema. Dato che lo stato futuro del sistema non è più predicibile a partire da quello presente, il fluido non sarà più incomprimibile (*b*); il suo comportamento assomiglierà più a quello di una nuvola di fumo che si espande fino a riempire tutto lo spazio di fase accessibile.

L'« universo giocattolo » consiste in una linea retta, infinita nei due sensi, e divisa in settori che possono essere occupati (*rettangoli scuri*) o vuoti (*rettangoli chiari*). Se la distribuzione dei rettangoli segue certe proprietà statistiche si può dimostrare che questo universo non contiene informazione microscopica. Per esempio, le proprietà specifiche di particolari successioni di rettangoli non hanno senso. Tali proprietà non possono distinguere una rappresentazione dell'universo da un'altra (*a*, *b*), dato che ogni successione finita esiste sicuramente in tutte le rappresentazioni infinite. Una particolare successione non può nemmeno designare un'unica posizione in una singola rappresentazione, dato che la stessa successione deve sicuramente ripetersi altrove (*c*). Questo ragionamento si può estendere all'universo reale, il quale soddisfa le condizioni statistiche richieste.

mazione macroscopica negativa. Infatti una definizione di questo tipo conduce immediatamente all'equazione presentata precedentemente: $H+I=H_{max}=I_{max}$. Interpretiamo ora H come entropia termodinamica e I come informazione macroscopica. L'entropia è allora sempre positiva o uguale a zero, e se il massimo dell'entropia resta costante, come deve accadere in un sistema chiuso, allora l'entropia deve aumentare col diminuire dell'informazione macroscopica. Abbiamo così ricondotto l'origine della freccia termodinamica del tempo a una proprietà degli stati iniziali dei sistemi chiusi. L'entropia di un sistema chiuso aumenterà solo se nel sistema è presente all'inizio l'informazione macroscopica mentre l'informazione microscopica è assente.

Le perturbazioni casuali

Queste condizioni iniziali speciali possono fornire una spiegazione della freccia termodinamica del tempo, ma non si tratta di una spiegazione molto soddisfacente. Perché in natura queste particolari condizioni iniziali sono regolarmente soddisfatte? L'informazione microscopica sembra abbastanza semplice da produrre. Perché compare solo negli stati finali dei sistemi naturali e mai in quelli iniziali? Che cosa significa affermare che l'informazione microscopica è assente da un certo stato? È sempre possibile acquisire tale informazione spendendo una quantità sufficiente di energia. Infine, che significato si può attribuire alla distinzione tra il livello macroscopico e quello microscopico di descrizione? Una risposta plausibile a questi problemi è stata data nel 1912 dal matematico francese Émile Borel. In tempi recenti i risultati di Borel sono stati riscoperti ed elaborati da John M. Blatt, Peter G. Bergmann, Joel L. Lebowitz e Philip Morrison.

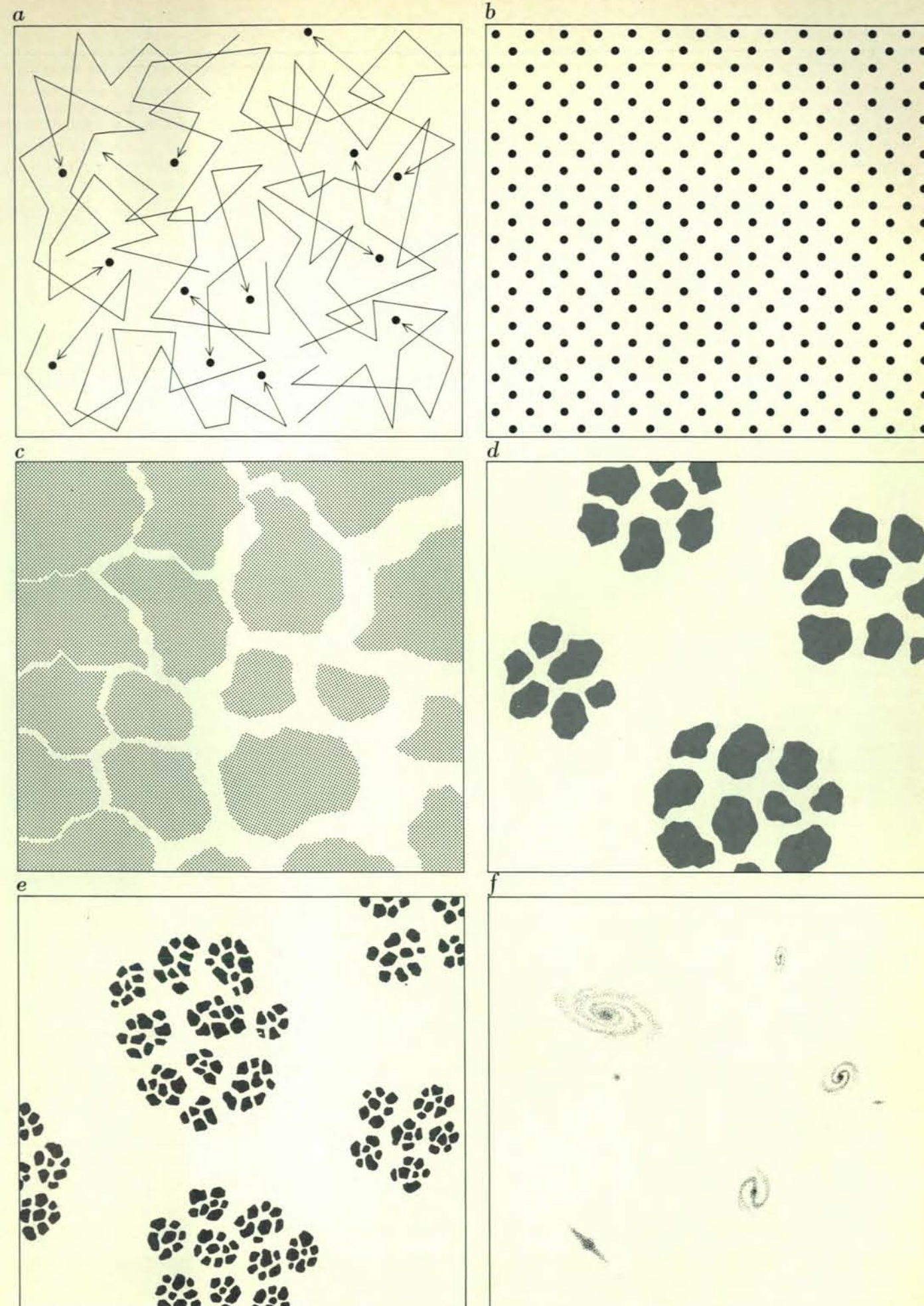
La nostra conclusione che l'informazione microscopica di un sistema aumenta col diminuire di quella macroscopica è valida solo per sistemi chiusi, cioè per sistemi che non comunicano con ciò che li circonda. Borel ha dimostrato che nessun sistema fisico finito può considerarsi chiuso. Si consideri per esempio la stanza in cui ha avuto luogo l'esperimento mentale della diffusione del profumo. Anche se la stanza non ha né porte né finestre, anche se le pareti sono isolate e molto spesse, il sistema di molecole non può essere isolato dal resto dell'universo. Le molecole di profumo e di aria entrano in collisione con le pareti della stanza, che sono in contatto col mondo esterno. Inoltre, fatto ancora più importante, è impossibile in linea di principio proteggere le molecole da interazioni gravitazionali con masse di materia anche lontane. Gli effetti di queste interazioni sono piccolissimi ma non trascurabili: Borel ha calcolato che il mutamento di potenziale gravitazionale prodotto dallo spostamento di un centimetro di un grammo di materia posto alla distanza di Sirio può alterare sostanzialmente nel corso di un microsecondo lo stato microscopico di un volume macroscopico di gas.

L'inevitabile interazione di un sistema chiuso solo di nome con il resto dell'universo ha l'effetto di una piccola perturbazione casuale che distrugge le correlazioni tra le velocità delle particelle. La perturbazione distrugge quindi l'informazione microscopica e ricrea di continuo le condizioni iniziali necessarie per garantire la degradazione dell'informazione macroscopica e la crescita dell'entropia termodinamica. Dato che non è più possibile considerare il sistema come isolato, la sua storia dinamica non sarà più completamente determinata. Il fluido probabilistico nello spazio di fase non sarà più incomprimibile, ma si espanderà come una nuvola di fumo fino a riempire tutto l'ipervolume disponibile (*si veda la figura di pag. 34*). Possiamo quindi concludere che nel mondo reale l'informazione macroscopica si trasforma in informazione microscopica, ma l'informazione microscopica viene distrutta da perturbazioni casuali.

Il principio cosmologico

Il ragionamento di Borel dipende dalla supposta casualità delle interazioni di sistemi nominalmente chiusi con il resto dell'universo. Se si conoscessero le posizioni e le velocità di tutte le particelle perturbanti si potrebbe estendere la definizione di sistema chiuso fino a includere le particelle perturbanti. Tuttavia anche questo sistema più grande sarebbe soggetto a perturbazioni provenienti dall'esterno. In ultima analisi saremmo costretti a includere nella nostra descrizione l'intero universo. Data una descrizione microscopica completa dell'universo (con tutte le limitazioni imposte dal principio d'indeterminazione) non ci sarebbe più nessuna distinzione qualitativa tra le due direzioni del tempo, perché una descrizione di questo tipo sarebbe simmetrica rispetto al tempo. Tuttavia una descrizione del genere è possibile, anche solo in linea di principio?

Ogni sistema fisico finito ammette una descrizione microscopica completa contenente una quantità finita d'informazione, e quindi si potrebbe pensare che anche l'universo nel suo complesso possa essere descritto compiutamente. Se la quantità d'informazione richiesta da una descrizione siffatta sia finita o no dipende solo dal volume finito o infinito dell'universo. (La cosmologia relativistica ammette entrambe le possibilità.) Tuttavia l'universo gode di certe proprietà specifiche che i suoi sottosistemi non condividono. In particolare ogni sottosistema finito dell'universo è limitato, mentre l'universo, finito o infinito che sia, si suppone illimitato. Inoltre esso sembra conformarsi a quello che chiamo principio cosmologico forte, secondo il quale nessuna proprietà statistica dell'universo definisce una posizione o direzione privilegiata nello spazio. Il principio cosmologico (normale), così denominato da Albert Einstein nel 1916, stabilisce che la distribuzione spaziale della materia e del movimento nell'universo è omogenea e isotropa, a parte irregolarità locali. La versione forte stabilisce inoltre che le irregolarità locali stesse devono essere omogenee e isotrope. Tale principio è strettamente connesso nelle sue conseguenze con la nostra ricerca sull'origine della freccia termodinamica del tempo. Dimostrerò che il principio cosmologico forte implica che l'informazione microscopica riguardo all'universo sia oggettivamente assente, nel senso che non è acquisibile o specificabile. Questa limitazione della nostra conoscenza rappresenta una indeterminazione di tipo cosmico, connessa ma diversa dall'indeterminazione richiesta dal principio d'indeterminazione di Heisenberg. Si tratta di una proprietà dell'universo nel suo complesso, non goduta però da sottosistemi limitati, per i quali l'informazione mi-

L'evoluzione dell'universo rappresenta una crescita dell'informazione macroscopica. In un modello progettato dall'autore e dai suoi collaboratori si suppone che lo stato iniziale sia privo d'informazione e di struttura. Nel periodo immediatamente successivo al « big-bang » (*a*) l'universo è in uno stato di equilibrio termodinamico mantenuto dalla rapida interazione tra particelle e dalla radiazione. Dopo un'espansione di circa 15 minuti l'universo si cristallizza, o congela, in una lega di idrogeno metallico ed elio (*b*). Poiché l'espansione cosmica continua questo universo solidificato si spezza in frammenti aventi più o meno la massa di un pianeta (*c*), che formano un « gas » nel senso che interagiscono tra loro frequentemente e casualmente come le molecole di un gas. Nel gas planetario possono verificarsi fluttuazioni di densità (*d*) per la coagulazione di gruppi di frammenti, fluttuazioni che diventano sempre più grandi col crescere delle dimensioni dei gruppi che si aggregano (*e*). Alla fine si costituisce una gerarchia di strutture, corrispondente alle stelle, galassie, gruppi di galassie che oggi vediamo (*f*).

croscopica si può specificare o acquisire liberamente.

È possibile illustrare la nozione di indeterminazione cosmica considerando un « universo giocattolo » costituito da particelle puntiformi distribuite casualmente ma con densità media uniforme lungo una linea retta infinita. Le proprietà statistiche di questo universo unidimensionale possono essere valutate con la massima precisione. Per esempio, possiamo valutare il numero medio di punti per unità di lunghezza con la massima precisione facendo la media su segmenti di retta sempre più lunghi. È possibile specificare qualche proprietà non statistica, o microscopica, di questo universo giocattolo? Che cosa costituirebbe una proprietà microscopica? Supponiamo di avere due rappresentazioni dell'universo in questione, identiche riguardo alle loro proprietà statistiche. Stabilisco che definire una proprietà microscopica significa trovare qualche modo di distinguere le due rappresentazioni, dato che l'unico tipo di informazione su cui si potrebbe fondare una tale distinzione è di tipo non statistico, e quindi microscopica.

Allo scopo di rappresentare gli effetti del principio d'indeterminazione dobbiamo dividere il nostro universo unidimensionale in celle di uguale lunghezza, dove la lunghezza rappresenta la precisione con cui si può specificare la posizione di una singola particella. Se allora specifichiamo il numero di particelle che occupa ogni singola cella, l'universo giocattolo viene a essere rappresentato da una successione di « numeri di occupazione » infinita e aperta da entrambi i lati. L'informazione microscopica è ora definita come un'informazione che ci potrebbe permettere di distinguere due successioni di numeri di occupazione dotate delle stesse proprietà statistiche (o macroscopiche). Per stabilire che due successioni sono differenti potremmo cercare di metterle in corrispondenza, cella a cella, per l'intera lunghezza. Tuttavia, dato che nessuna delle due successioni ha un inizio o una fine o un qualsiasi altro punto in qualche modo privilegiato, esistono infiniti modi di confrontarle. In linea di principio è impossibile eseguire una serie infinita di operazioni e quindi non è questa la via attraverso cui si può dimostrare l'impossibilità di una corrispondenza tra le due successioni.

Si potrebbe invece cercare di dimostrare che le due successioni sono identiche. Innanzitutto dovremmo formare da una successione di numeri di occupazione una sottosuccessione di lunghezza qualsiasi, quindi potremmo cercare un'identica sottosuccessione nell'altra successione. In una successione infinita qualsiasi sottosuccessione di lunghezza finita si presenta un numero infinito di volte. La legge dei grandi numeri garantisce il successo della nostra ricerca in un numero finito di tentativi. Inoltre l'esito positivo non dipende in alcun modo dalla lunghezza della sottosuccessione scelta, purché sia finita. Le due successioni sono indistinguibili dal punto di vista operazionale perché, se non fosse così, dovrebbe essere possibile mostrare almeno una sottosuccessione di una successione che non possiede duplicato nell'altra successione. Possiamo quindi concludere che esiste solo una successione infinita di cifre con le proprietà statistiche che definiscono l'universo giocattolo. Due rappresentazioni dell'universo con le stesse proprietà statistiche sono indistinguibili. Dato che l'informazione microscopica è, per definizione, ciò che potrebbe distinguere le due successioni, dobbiamo concludere che essa è oggettivamente assente.

Questo ragionamento si può estendere facilmente a modelli infiniti dell'universo reale tridimensionale che soddisfino il principio cosmologico forte e la condizione aggiuntiva che la scala della struttura locale sia finita. La disposizione delle stelle e delle galassie visibili dalla Terra è talmente complessa e specifica che sembrerebbe definire la nostra posizione nell'universo con la stessa precisione con cui un'impronta digitale identifica una persona, ma non è così. In un universo infinito, statisticamente omogeneo e isotropo c'è la certezza che la stessa disposizione di stelle e galassie si presenti ripetutamente. Se il nostro universo soddisfa il principio cosmologico forte le sue proprietà significative sono tutte statistiche e il suo stato microscopico è completamente indeterminato. A partire dal tempo di Newton è stato implicito nel pensiero cosmologico che l'universo, in linea di principio, possa essere completamente descritto a livello microscopico. Ora possiamo vedere come ciò non sia affatto necessario. Se l'universo è sufficientemente simmetrico non vi è spazio per l'informazione microscopica.

L'origine dell'informazione macroscopica

Abbiamo visto come la freccia termodinamica del tempo tragga origine dall'assenza di informazione microscopica e dalla presenza di informazione macroscopica negli stati iniziali dei sistemi chiusi. Abbiamo trovato che in un universo che soddisfi il principio cosmologico forte l'informazione microscopica è oggettivamente assente. D'altra parte non siamo stati in grado di trovare ragioni che impedissero la mancanza anche dell'informazione macroscopica. La complessità dell'universo, come ci è svelata dall'astronomia, è davvero sconcertante. I sistemi isolati si avviano inevitabilmente verso uno stato indifferenziato di equilibrio termodinamico. Dato che anche l'universo è in un certo senso un sistema isolato, perché non ha ancora raggiunto l'equilibrio? Una risposta, che molti cosmologi sono pronti a fornire, è che di fatto l'universo tende a questo stato di equilibrio, ma che deve passare ancora del tempo prima che il processo sia compiuto. Fred Hoyle e J. V. Narlikar hanno scritto: « Nella cosmologia del « big-bang » l'universo deve partire con un forte squilibrio termodinamico e deve alla fine esaurirsi. » Dimostrerò che questo punto di vista è fondamentalmente scorretto. L'universo non va verso l'esaurimento e non è necessario che all'inizio della sua storia si sia verificato un forte squilibrio termodinamico; infatti lo stato iniziale può essere stato completamente privo sia di informazione macroscopica sia di informazione microscopica.

Supponiamo che in qualche momento vicino all'origine si a prevalso nell'universo un equilibrio termodinamico locale. L'entropia di ogni regione sarebbe stata allora la più grande possibile in base ai valori prevalenti di temperatura media e di densità. Con l'espandersi dell'universo da quello stato iniziale ipotetico, i valori locali di densità media e di temperatura si sarebbero modificati, e lo stesso avrebbe fatto l'entropia di ogni regione. Per mantenere l'entropia al suo valore massimo (e quindi per il mantenimento dell'equilibrio) la distribuzione dell'energia nella materia e nella radiazione deve variare, e lo stesso vale per la concentrazione dei vari tipi di particelle. I processi fisici attraverso cui si effettuano questi mutamenti procedono a un ritmo finito; se questi ritmi di « equilibrazione » sono molto più veloci di quelli dell'espansione cosmica, si manterrà localmente un approssimato equilibrio termodinamico, in caso contrario l'espansione darà origine a sensibili spostamenti dall'equilibrio a livello locale. Questi spostamenti rappresentano l'informazione macroscopica; la quantità d'informazione macroscopica generata dall'espansione è la differenza tra il valore attuale dell'entropia e il valore massimo teorico della medesima a livelli medi di temperatura e densità.

Questo ragionamento non dipende dall'ipotesi di una semplice espansione cosmica, ma dal ritmo finito a cui possono avvenire i mutamenti di densità e temperatura. Le stesse conclusioni si potrebbero trarre se l'universo si stesse contraendo da uno stato di equilibrio invece che espandendo: se il ritmo della contrazione fosse più veloce di quello di questi processi che mantengono l'equilibrio termodinamico, aumenterebbero sia l'informazione macroscopica sia l'entropia. Il nostro risultato non fissa quindi la direzione della freccia cosmologica rispetto a quella termodinamica, ma si limita a stabilire che l'informazione macroscopica e l'entropia sono generate dall'evolversi dell'universo da un ipotetico stato iniziale di equilibrio termodinamico locale.

È ragionevole supporre che l'universo sia stato (o sarà mai) in uno stato di equilibrio termodinamico locale? Per rispondere a questa domanda dobbiamo confrontare i ritmi dei processi di equilibrazione (quelli che generano entropia) e quelli dell'espansione o contrazione cosmica. Nessuno dei due è costante. Spostandoci indietro nel tempo verso l'origine il ritmo d'espansione aumenta, e all'origine del tempo il ritmo di espansione è infinito. Tuttavia anche i ritmi di equilibrazione aumentano man mano che ci spostiamo indietro nel tempo, dato che gli scontri tra particelle diventano sempre più frequenti con l'aumento della densità e della temperatura. Di fatto, nel periodo immediatamente successivo all'origine, i ritmi dei processi di equilibrazione sono molto più alti di quello dell'espansione cosmica. Ne deriva che il « big-bang » si presenta come un processo estremamente dolce, in cui i processi di equilibrazione locali tengono facilmente il passo delle mutazioni delle condizioni macroscopiche di temperatura e densità che avvengono nella prima frazione di un microsecondo. L'equilibrio termodinamico locale si può supporre solo per questa brevissima fase iniziale dell'evoluzione dell'universo, ma da questa assunzione deriva che l'espansione dell'universo ha generato sia l'informazione macroscopica sia l'entropia. Quindi la freccia cosmologica, quella storica e quella termodinamica si presentano come conseguenze del principio cosmologico forte e dell'assunzione che all'origine, o in un momento a essa molto vicino, sia prevalso uno stato di equilibrio termodinamico locale. È degno di nota che nessuna di queste assunzioni si riferisca direttamente al tempo o a processi temporali.

Rimane una domanda aperta, se si considera plausibile questa teoria. L'espansione cosmica genera i particolari tipi di informazione macroscopica che caratterizzano l'universo oggi? È possibile che parte dell'informazione sia stata presente fino dall'inizio, forse sotto forma di fluttuazioni di densità. Non è possibile ancora rispondere con sicurezza alla domanda, ma è importante notare che da un punto di vista teorico non è necessario che nello stadio iniziale fosse presente qualche tipo di struttura. Insieme ai miei collaboratori ho sviluppato un modello dell'evoluzione in cui si parte da una situazione di equilibrio termodinamico a temperatura zero (*si veda la figura di pagina 37*). È quindi possibile che l'universo, con tutta la sua ricchezza e la differenziazione delle sue parti, si sia sviluppato da uno stato completamente privo di informazione e di struttura. Se postuliamo l'esistenza di uno stato primordiale di questo tipo possiamo anche fare a meno di formulare separatamente il principio cosmologico forte. L'omogeneità statistica e l'isotropia dell'universo derivano dall'invarianza di tutte le leggi fisiche note rispetto a rotazioni e a traslazioni spaziali.

Novità e determinismo

Abbiamo ora ricondotto a un'origine comune la freccia termodinamica e quella storica, vale a dire allo stato iniziale dell'universo. In tale stato l'informazione microscopica è assente e quella macroscopica è assente o minima. L'espansione a partire da questo stato ha prodotto tanto l'entropia quanto la struttura macroscopica. L'informazione microscopica, d'altra parte, non è presente nei sistemi astronomici di formazione recente, ed è questo il motivo per cui essi e i loro sottosistemi possiedono una freccia termodinamica.

La concezione di un mondo che si evolve nel tempo è radicalmente differente da quella che ha dominato nella fisica e nell'astronomia a partire dai tempi di Newton, una concezione che trova la sua espressione classica nelle opere di Pierre Simon de Laplace.

In questa prospettiva non c'è niente che corrisponda al passaggio del tempo, il passato e il futuro coesistono come termini uguali, come due semirette staccate da un punto che divide una retta. Tuttavia, se la teoria che ho presentato è corretta, nemmeno il sommo calcolatore, l'universo, contiene informazione sufficiente per specificare in modo completo i suoi stati futuri. Il momento presente contiene sempre un elemento di genuina novità e il futuro non è mai completamente predicibile. Poiché i processi biologici generano anch'essi informazione e la coscienza ci permette di sperimentare in modo diretto questi processi, la percezione intuitiva del mondo come qualcosa che si espande nel tempo coglie una delle caratteristiche più profonde dell'universo.

TEMI METAMAGICI

di Douglas R. Hofstadter

Strani attrattori: schemi matematici collocati fra l'ordine e il caos

«Non puoi sapere quanto felice
io sia d'averti incontrato.
Ho una strana attrazione per te.»
—COLE PORTER
«It's All Right with Me»

Alcuni mesi fa, mentre passeggiavo con un amico per i corridoi del dipartimento di fisica dell'Università di Chicago, mi cadde l'occhio su un manifesto che annunciava un simposio internazionale dal titolo «Strani attrattori». Non potei fare a meno di essere attratto da questo strano termine e chiesi al mio amico di cosa si trattasse. Mi rispose che era un argomento di grande attualità e, dalla descrizione che me ne fece, mi apparve veramente incantevole e pieno di mistero.

Riuscii a capire che l'idea di fondo stava nell'analisi di quelli che si potrebbero chiamare cicli di retroazione: situazioni in cui il risultato della computazione si può reintrodurre come un nuovo argomento, allo stesso modo in cui i suoni che escono da un altoparlante possono rientrare nel microfono c nuovamente uscire. Sembrava che dal più semplice di questi cicli potessero emergere sia strutture stabili sia strutture caotiche (se questa non è una contraddizione in termini!): la differenza stava solo nel valore di un unico parametro. L'ordine o non ordine del sistema ciclico dipendeva da piccolissime variazioni nel parametro e questa immagine dell'ordine che svanisce dolcemente nel caos, della struttura che si dissolve progressivamente nella casualità, mi sembrava pieno di fascino.

Sembrava anche che fossero di recente venute alla luce delle inaspettate caratteristiche «universali» della transizione nel caos, caratteristiche che dipendevano unicamente dalla presenza del *feedback* che erano virtualmente insensibili ad altri dettagli del sistema. Questa generalità era importante perché un modello matematico che evidenzi un approccio graduale al comportamento caotico potrebbe costituire la chiave per la comprensione della comparsa della turbolenza in tutti i tipi di sistemi fisici. La turbolenza, al contrario della maggior parte dei fenomeni fisici di cui si è raggiunta la comprensione, è un fenomeno non lineare: due soluzioni alle equazioni di turbolenza non conducono insieme a una nuova soluzione. Lo sviluppo della matematica non lineare è molto inferiore a quello della matematica lineare, e questo è il motivo per cui è a lungo sfuggita ai fisici una descrizione matematica della turbolenza, descrizione che, per ovvi motivi, sarebbe di una importanza fondamentale.

Quando in seguito cominciai a leggere qualcosa a proposito di queste nozioni, scoprii che esse erano scaturite simultaneamente da molte discipline. Gli studiosi di matematica pura avevano cominciato a occuparsi dell'interazione dei sistemi non lineari utilizzando i calcolatori. Gli studiosi di metereologia teorica e i genetisti della popolazione, così come i fisici teorici impegnati in campi diversi quali i fluidi, i laser e le orbite planetarie, erano indipendentemente arrivati a modelli matematici non lineari di tipo analogo, caratterizzati da cicli di retroazione immersi nel caos; studiando le proprietà di questi modelli, ciascun gruppo trovava particolarità che gli altri non avevano trovato. Inoltre, non solo i teorici, ma anche gli sperimentatori operanti all'interno di queste discipline così diverse avevano compiuto simultaneamente osservazioni su fenomeni caotici che denotavano una struttura di fondo comune. Mi accorsi presto che la semplicità delle idee base conferisce loro un'eleganza che rivaleggia con quella di alcuni tratti della matematica classica; c'è effettivamente, in alcuni aspetti di questo lavoro, un sapore di XVIII o XIX secolo che trovo particolarmente piacevole in quest'epoca di impressionante astrazione.

Esempi di funzioni non monotone possono generare un picco aguzzo (a sinistra) e una parabola (a destra)

Probabilmente, la ragione principale per cui queste idee vengono alla luce solo ora è che le modalità di ricerca sono del tutto moderne: si tratta di un tipo di matematica sperimentale in cui il calcolatore svolge un ruolo quale quello della nave per Magellano, del telescopio per gli astronomi e dell'acceleratore per i fisici. Come le navi, i telescopi e gli acceleratori devono essere sempre più grandi, più potenti e più costosi per sondare sempre più in profondo la natura, così ci vorrebbero calcolatori sempre più grossi, più veloci e affidabili per esplorare le regioni più lontane dello spazio matematico. Come ci fu un'epoca d'oro per le esplorazioni navali e per le scoperte fatte con i telescopi e gli acceleratori, epoca caratterizzata da un elevato rapporto tra nuovi segreti svelati e capitali stanziati, così ci si dovrebbe attendere nella matematica sperimentale un'epoca d'oro di questi modelli del caos. Forse quest'epoca si è già aperta, o forse si sta aprendo proprio ora; e forse in seguito assisteremo a un fiorire di lavori teorici fondati su queste scoperte sperimentali.

Si tratta comunque di un modo curioso e incantevole di far matematica, capace di costruire direttamente nella mente efficaci immagini e intuizioni visive. Superando il tradizionale schema matematico teorema-dimostrazione-teorema-dimostrazione, si sfrutta la potenza dei calcolatori per arrivare rapidamente a osservazioni empiriche, a scoperte che si rinforzano le une con le altre; si forma così una rete talmente ricca e coerente di idee che alla lunga può risultare più facile trovare le necessarie dimostrazioni, data la completezza della rilevazione a cui è stato sottoposto in precedenza il territorio concettuale. Uno dei più decisi sostenitori di questo modo di far matematica è stato Stanislaw M. Ulam, il quale, in epoca ancora pionieristica, sottoponeva ai calcolatori problemi di iterazione non lineare, oltre a problemi relativi a molte altre branche della matematica. Molte delle idee delineate in questo articolo sono tratte dai primi studi effettuati da Ulam insieme a Paul Stein.

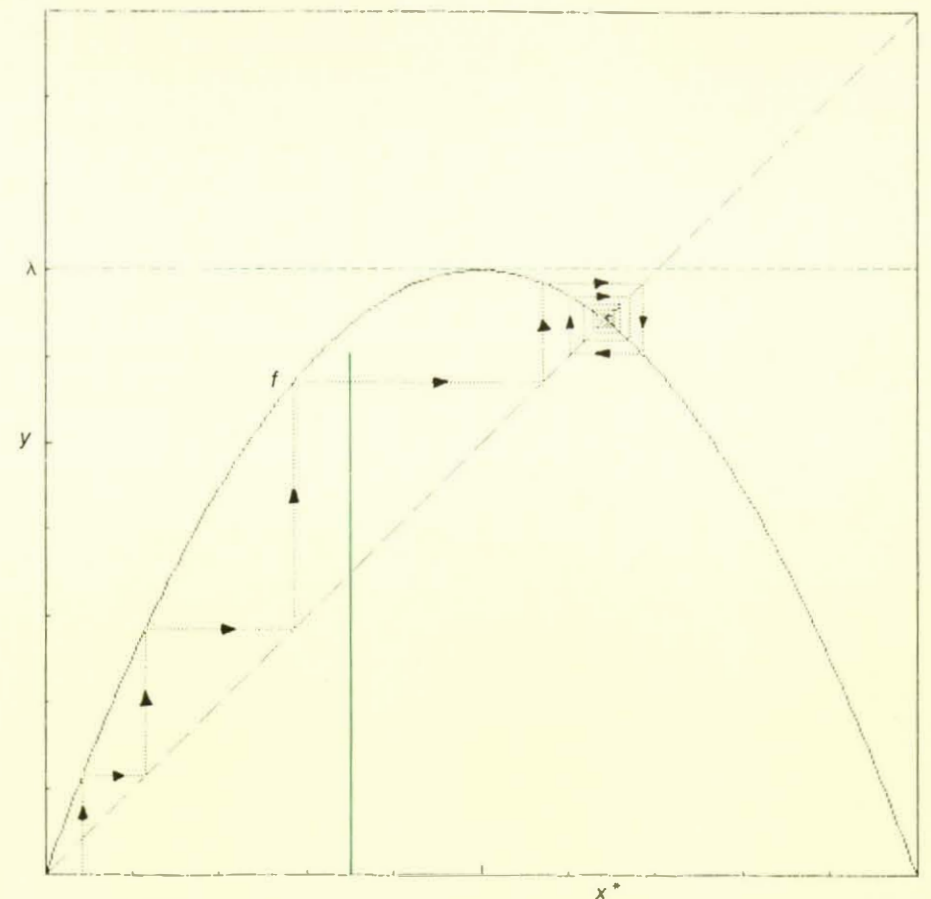
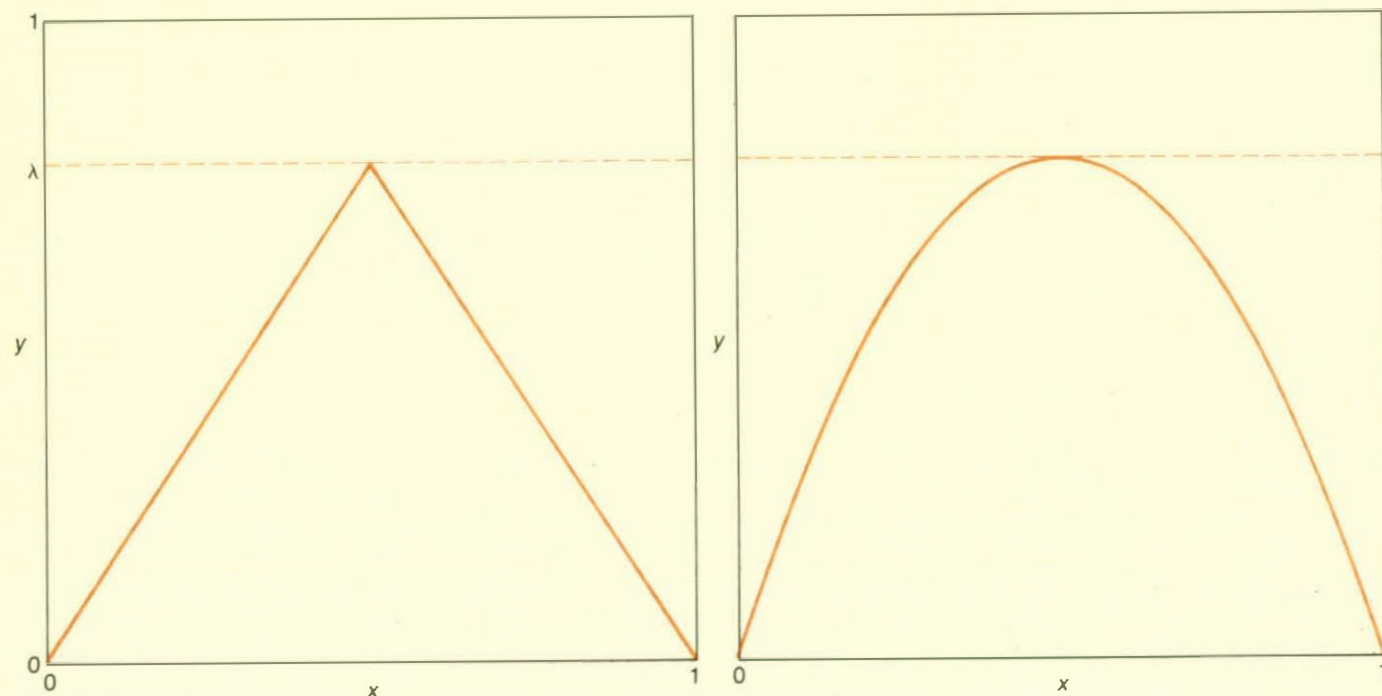
Ma vediamo ora di arrivare alla nozione di strano attrattore, partendo dal concetto base di attrattore. Tutto il discorso si fonda su un'unica idea: l'iterazione di una funzione matematica a valori reali, cioè il comportamento della sequenza di valori $x, f(x), f(f(x)), f(f(f(x))), \ldots$, dove f è una certa funzione a cui siamo interessati. Il valore iniziale di x può essere chiamato il «seme». L'idea è quella di continuare a reintrodurre in f il valore di f e vedere se emerge una qualche struttura.

Ecco un problema interessante e non troppo difficile a proposito dell'iterazione di una funzione: si può inventare una funzione p con la proprietà che per un qualsiasi valore reale di x, anche $p(x)$ e in cui $p(p(x))$ sia uguale a $-x$? Il problema è complicato dalla condizione che $p(x)$ sia reale, altrimenti la funzione $p(x)$ (d

λ

y

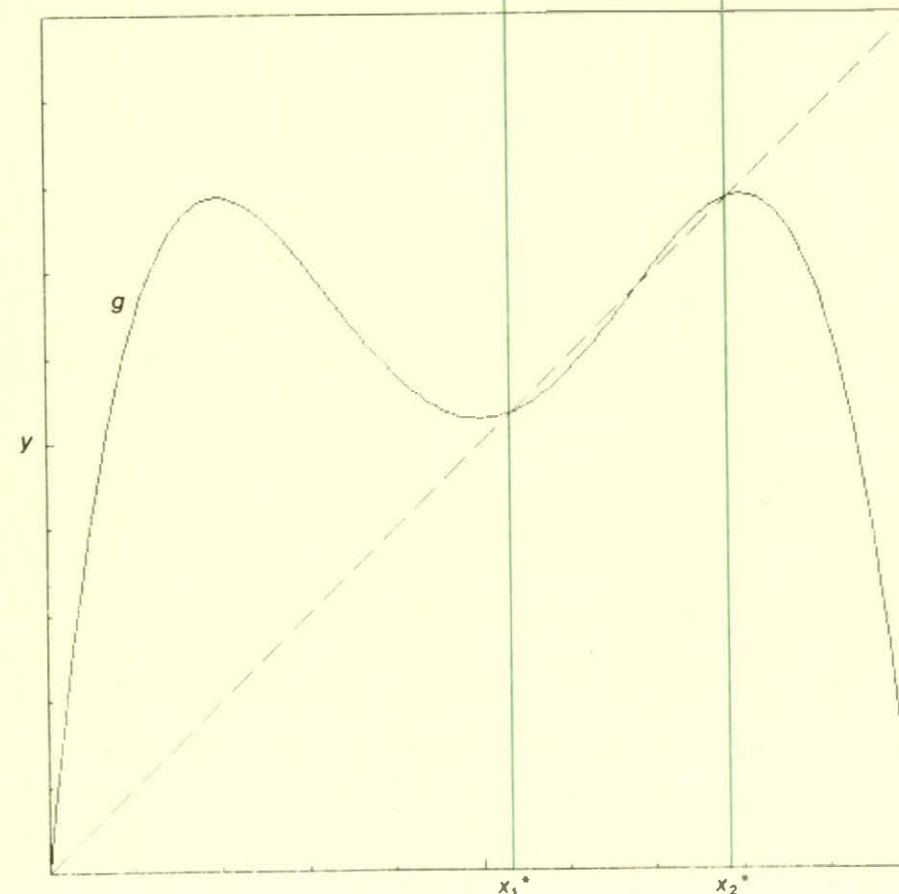
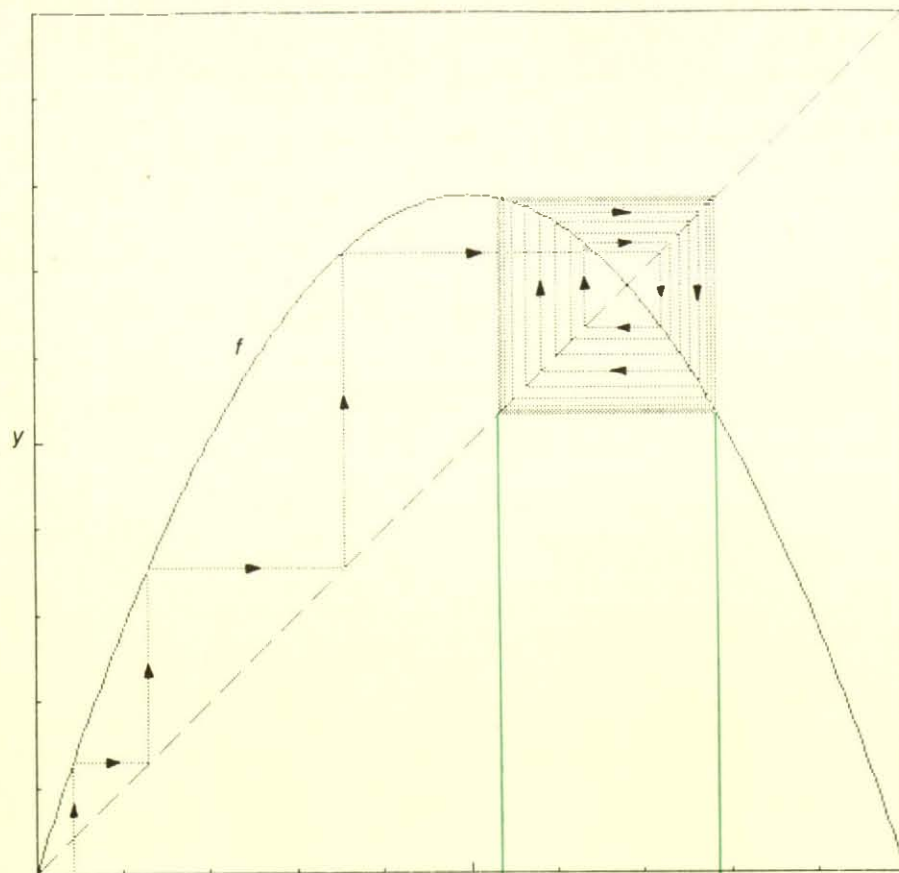
f

x^*

Regolando la «manopola λ» al livello 0,7 si ottiene questo grafo per la funzione f(x)

be al caso nostro. In realtà, la questione equivale alla ricerca di una «radice quadrata del segno meno» a valore reale. Un problema connesso è quello di trovare una funzione a valori reali q con la proprietà che $q(q(x)) = 1/x$ per tutti gli x diversi da zero. Si noti che comunque si costruiscano p e q, entrambe avranno la proprietà che, dato un qualsiasi seme, l'iterazione ripetuta crea un ciclo di lunghezza quattro.

Più in generale, quali tipi di funzioni è probabile che mostrino, se ripetutamente iterate, un interessante comportamento ciclico o quasi ciclico? Nulla del genere avviene eon semplici funzioni come $3x$ o x^3. L'n-esima interazione di $3x$, per esempio, è $3 \times 3 \times 3 \times \ldots \times 3 \times x$, dove il 3 è ripetuto n volte - cioè $3^n x$ - e l'n -esima interazione di x^3 non è che $(((x^3)^3)^3 \ldots)^3$, dove il 3 è ripetuto n volte, che assomma a x^{3^n}. Non c'è nulla di ciclico qui: i valori continuano semplicemente a crescere. Per rovesciare questa tendenza è necessario avere una funzione con una specie di scambio, un piccolo zig-zag, una curvatura. Per dirla in termini più tecnici, bisogna avere una funzione *non monotona*: una funzione il cui grafo sia ondulato, cioè che si muova in una direzione, diciamo verso l'alto, e poi curvi verso l'altra direzione, diciamo verso il basso.

Sulla sinistra della figura della pagina a fronte vediamo un dente di sega con una punta acuminata al vertice, mentre a destra abbiamo un arco parabolico che si incurva dolcemente. Entrambe le curve nascono dall'origine, raggiungono un picco indicato da λ e poi scendono per atterrare sul lato opposto dell'intervallo. Naturalmente ci sono un'infinità più che numerabile di forme che salgono all'altezza λ e poi ridiscendono, ma queste sono tra le più semplici. Delle due, la parabola è forse la più semplice, o almeno quella di maggior interesse matematico; la sua equazione è $y = 4\lambda x(1-x)$, con λ non superiore a 1.

Ammetteremo in entrata - come valori di x - solo numeri compresi tra 0 e 1. Come mostra il grafo, per ogni x in quell'intervallo, l'uscita - y - è sempre compresa tra 0 e λ. Il valore dell'output, quindi, può essere sempre reintrodotto come entrata nella funzione, ciò che ci assicura che l'iterazione sarà sempre possibile. Quando si ripete l'iterazione di una funzione «ondulata» come questa, i successivi valori y che si producono a volte saliranno e a volte scenderanno, sempre oscillando, naturalmente, tra 0 e λ. L'ondulazione del grafo garantisce, come vedremo, interessanti effetti quando la funzione viene iterata.

Risulta che le spettacolari differenze nel grado di regolarità a cui ho accennato prima sono dovute alla variazione nella posizione di ciò che potremmo chiamare la «protuberanza λ». A seconda del valore a cui si assesta la protuberanza, la funzione descrive un'incredibile varietà di «orbite», cioè di sequenze $x, f(x), f($

In alto, una spirale verso un ciclo-2 stabile; in basso, gli elementi del ciclo x_1^* *e* x_2^*

e così via. In particolare, per λ al di sotto di un certo valore critico (λ_c = 0, 892486417967...) le orbite sono tutte regolari e strutturate (anche se esistono vari *gradi* di strutturazione; di solito, più basso è λ, più semplice è la struttura dell'orbita), ma per λ pari o superiore a questo valore critico, tenetevi stretti! Una sequenza di valori essenzialmente caotica verrà tracciata dai valori x, $f(x)$, $f(f(x))$

vo scelto. Nel caso della parabola, sembra che il primo ad accorgersi del ruolo critico svolto dalla variazione della protuberanza λ sia stato P. J. Myrberg agli inizi degli anni sessanta, ma il suo lavoro venne pubblicato su un'oscura rivista e non ricevette molta attenzione. Qualcosa come 10 anni più tardi, Nicholas C. Metropolis, Paul Stein e Myron Stein riscoprirono l'importanza della protuberanza λ, non solo per la parabola ma per molte funzioni. In effetti, essi scoprirono che finché erano in gioco certe proprietà topologiche, non aveva importanza la funzione, ma solo il valore di λ. A questo fatto è stato dato il nome di «universalità strutturale».

Per capire come funzioni questa non intuitiva dipendenza dalla posizione della protuberanza λ bisogna arrivare a visualizzare il processo di iterazione di $f(x)$. La cosa è semplice. Supponiamo che il valore di λ sia 0,7. Il grafo di $f(x)$ è quello che si vede nella figura della pagina precedente, insieme alla linea spezzata a 45 gradi $y = x$. (Questo grafo e molti altri di questo articolo sono stati prodotti su un piccolo calcolatore da Mitchell J. Feigenbaum del Los Alamos National Laboratory.)

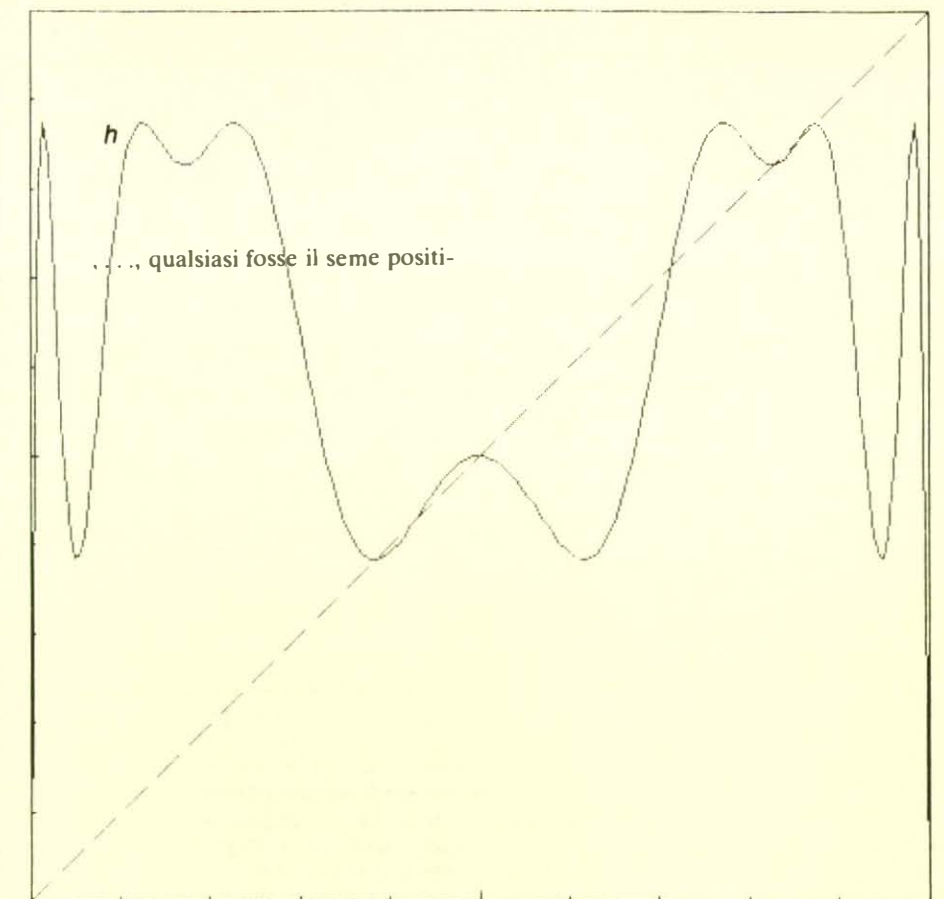
Consideriamo i due valori di x in cui si intersecano la linea a 45 gradi e la curva: essi sono $x = 0$ e $x = 9/14 \approx 0,643$. Indichiamo il valore diverso da zero con x^*. Allora, per costruzione, $f(x^*)$ è uguale a x^* e la ripetuta iterazione di f a questo valore x porterà a un ciclo infinito. Lo stesso avviene se si inizia a iterare a $x = 0$: si rimane impigliati in un ciclo senza fine. C'è però una differenza significativa tra questi due «punti fissi» di f e la cosa risulta meglio se si prende qualche altro valore iniziale di x, ad esempio uno vicino a 0,04, come si vede nella stessa figura. Chiamiamo x_0 questo valore iniziale di x. C'è un elegante modo grafico per generare l'orbita di qualsiasi seme x_0. Una linea verticale in x_0 incontra la curva all'altezza $y_0 = f(x_0)$. Per iterare f bisogna disegnare una nuova linea verticale passante per il nuovo valore di x uguale a questo valore di y. Ed è qui che viene utile la linea a 45 gradi $y = x$: rimanendo all'altezza y_0, ci spostiamo orizzontalmente fino a raggiungere la linea a 45 gradi; dato che lungo questa linea y è uguale a x, sia x sia y sono uguali a y_0. Chiamiamo x_1 questo nuovo valore di x. Tracciamo ora una seconda linea verticale, la quale incontrerà la curva all'altezza $y_1 = f(x_1) = f(y_0) = f(f(x_0))$. Quindi si ripete il procedimento.

In breve, basta una semplice regola per realizzare l'iterazione: (1) spostarsi verticalmente fino a raggiungere la curva e (2) poi spostarsi orizzontalmente fino a raggiungere la linea diagonale. Continuare a ripetere i passi (1) e (2).

Nella figura si vede anche il risultato di questo procedimento con seme x_0 = 0,04: un vivace girotondo intorno al punto con x^* come ascissa e ordinata. A poco a poco ci si avvicina a x^*, che è quindi un tipo peculiare di punto fisso perché *attrae* i valori iterati di $f(x)$. È questo l'esempio più semplice di attrattore: ogni possibile seme (tranne 0) è trascinato, mediante l'iterazione di f, a questo valore stabile di x. Per questa ragione x^* è chiamato punto fisso stabile. Al contrario, 0 è un punto fisso repulsivo o non stabile, dato che l'orbita di qualsiasi valore iniziale x, per quanto vicino a 0, si allontanerà da 0 per muovere verso x^*. Si noti che a volte le iterazioni di f sorpasseranno x^* e a volte cadranno bruscamente, ma inesorabilmente si avvicinano sempre più a x^*, puntando su di esso come rondini che tornano al nido.

Qual è la ragione di questa radicale differenza qualitativa tra i due punti fissi (0 e x^*) di f? Basta dare un'occhiata alla figura per rendersi conto che la ragione sta nel fatto che in 0 la curva si inclina troppo ripidamente. In particolare, la curvatura è superiore in quel punto a 45 gradi ed è l'inclinazione locale della curva a determinare di quanto ci si deve muovere orizzontalmente ogni volta che si itera f. Se la curva ha un'inclinazione superiore a 45 gradi (sia che vada verso l'alto sia che vada verso il basso), essa tende a spingervi sempre più lontano dal vostro punto di partenza man mano che si itera secondo le regole (1) e (2). Quindi il criterio per la stabilità di un punto fisso è che la curvatura nel punto sia minore di 45 gradi. Ora, è questo il caso di x^* quando λ è a 0,7; infatti lì la curvatura è di circa 41 gradi, mentre in 0 è molto maggiore di 45 gradi.

Cosa avviene se aumentiamo λ? La posizione di x^* (x^* è per definizione il punto in cui si intersecano la linea f e la linea $y = x$) cambia e anche la pendenza di f in x^* aumenta. Cosa succede quando la pendenza raggiunge 45 gradi o li supera? Questo avviene quando il valore di λ è pari a 3/4 e chiameremo Λ_1 questo particolare valore di λ. Guardiamo una figura in cui il valore di λ è leggermente superiore, vale a dire λ = 0,785 *(si veda la figura della pagina a fronte)*.

Cosa avviene, invece, se iniziamo con un seme scelto a caso, diciamo ancora x = 0,04? Nella metà superiore della figura si vede l'orbita risultante. Come si può vedere avviene una cosa molto bella: dapprima i valori salgono nelle vicinanze di x^* (ora un punto fisso instabile di f), ma poi si allontanano gradatamente a spirale e si assestano in una specie di danza quadrata convergente su due particolari valori x_1^* e x_2^*. Questa elegante oscillazione è chiamata ciclo-2 e la coppia di valori di x che lo costituiscono (x_1^* e x_2^*) è ancora chiamata un attrattore, per precisione un attrattore di periodo due. Questo termine indica che il nostro ciclo-2 è stabile: con l'iterazione di f esso attrae i valori di x. Nella stessa danza finirà l'orbita di qualsiasi valore positivo del seme (tranne x^*

Il grafo di $h(x)$ *per* λ *pari a 0,87 è ancora più a gobbe di quello di* $g(x)$

stesso), cioè si avvicinerà asintoticamente al perfetto ciclo-2 composto dai punti x_1^* e x_2^*, anche se non lo raggiungerà mai esattamente. Dal punto di vista di un fisico, comunque, l'accuratezza dell'avvicinamento diviene presto così grande da poter ugualmente dire che le orbite sono state «catturate» dall'attrattore.

Per capire bene questo fatto si può guardare il grafo di una nuova funzione ricavata da quella vecchia. Consideriamo il grafo di $g(x) = f(f(x))$ che si vede nella metà inferiore della figura. Questo cammello a due gobbe è chiamato *l'iterata* di f. Osserviamo dapprima che qualsiasi punto fisso di f è anche un punto fisso di g, quindi 0 e x^* saranno punti fissi di g. Ma si osserva poi che essendo $f(x_1^*)$ uguale a x_2^*, e per converso $f(x_2^*)$ uguale a x_1^*, g avrà due *nuovi* punti fissi: $g(x_1^*) = x_1^*$ e $g(x_2^*) = x_2^*$. Nella rappresentazione grafica è facile individuare x_1^* e x_2^*: sono i punti in cui si intersecano la linea a 45 gradi e il grafo a due gobbe di $g(x)$. Ci sono quattro di questi punti (0 e x^* sono gli altri due). Come abbiamo visto, il criterio per la stabilità di qualsiasi punto fisso sotto iterazione è che la pendenza in quel punto sia minore di 45 gradi. Qui abbiamo a che fare con punti fissi di g e quindi con la pendenza di g (distinta da quella di f). Infatti, nella stessa figura si può chiaramente vedere che in 0 e in x^* g ha una pendenza maggiore di 45 gradi, mentre sia in x_1^* sia in x_2^* la pendenza di g è minore di 45 gradi. In effetti, cosa rimarchevole, non solo entrambi i valori della pendenza sono minori di 45 gradi, ma anche, come risulta da un semplice calcolo, uguali (o «schiavi» l'uno dell'altro, come a volte si dice).

Abbiamo visto un attrattore di periodo uno convertito in un attrattore di periodo due dato un particolare valore di λ (λ = 3/4). Proprio a quel valore il punto fisso x^* si divide in due valori oscillanti, x_1^* e x_2^*. Naturalmente essi coincidono alla «nascita», ma con l'aumentare di λ essi si separano e si allontanano sempre più. Questo aumento di λ porta anche alla crescita della pendenza di g in questi due punti fissi stabili (di g) finché, per qualche valore di λ, raggiunge, come il suo progenitore f, il proprio punto di rottura (cioè le pendenze identiche in x_1^* e x_2^* superano i 45 gradi) ed entrambi i due punti d'attrazione perdono d'efficacia generando i propri cicli-2 locali. (In realtà i cicli sono cicli-2 solo finché si tratta di g; per f, i nuovi punti sono elementi di un attrattore di periodo quattro. Bisogna stare attenti a tenere bene a mente f e g!). Queste due divisioni avverranno esattamente nello stesso momento (il che equivale a dire per lo stesso valore di λ), dato che il valore della curvatura di g in x_1^* è schiavo del valore della pendenza in x_2^*. Questo valore di λ sarà chiamato Λ_2 ed è pari a 0,86237...

Ormai afferrato il principio, avrete già

probabilmente intuito che per qualche nuovo valore Λ_3 tutti e quattro i punti dell'attrattore di f si divideranno simultaneamente, dando luogo a un attrattore formato da otto punti; e che poi questo schema procede allo stesso modo, con continui raddoppi una volta raggiunti e sorpassati certi particolari valori di λ. Se questa è la vostra previsione, avete perfettamente ragione e il motivo di fondo è sempre lo stesso: le pendenze (identiche) in tutti i punti fissi stabili di qualche grafo raggiungono l'angolo critico di 45 gradi. Nel caso della prima divisione (in Λ_1) si trattava della pendenza della stessa f nell'unico punto x^*. La divisione successiva era dovuta alle pendenze nei due punti fissi stabili x_1^* e x_2^* di g, pendenze che raggiungono simultaneamente i 45 gradi. Analogamente, Λ_3 è quel valore di λ per il quale la pendenza di $h(x) = g(g(x)) = f(f(f(f(x))))$ tocca i 45 gradi contemporaneamente nei quattro punti fissi stabili di h. E così via. Nella figura della pagina precedente si vede l'aspetto tormentato di $h(x)$ per un valore di λ prossimo a 0,87.

Nella figura qui sotto si vede la posizione sull'asse delle x dei punti fissi stabili di f per Λ_1 - Λ_6 (ce ne sono 32, alcuni così ravvicinati tra loro da non poter essere distinti). I punti sono riprodotti nel momento in cui stanno per divenire instabili, come una cellula sul punto di dividersi. Si noti la rigorosa disposizione dei punti d'attrazione: guardando questa rappresentazione grafica della spaziatura tra gli elementi dei successivi attrattori di f a periodo raddoppiato, si può vedere che ogni linea può essere ricavata da quella sopra mediante uno schema geometrico ricorsivo per cui ogni punto è rimpiazzato da due punti «gemelli» sotto di esso. Ogni raggruppamento locale di punti è semplicemente una riduzione in scala dello schema globale (e anche un'inversione sinistra-destra nell'alternarsi dei gruppetti locali). Per esempio, nella linea in basso è stato evidenziato con il colore un gruppo locale di otto punti: si noti come il gruppo di punti sia una versione in miniatura dello schema globale due linee sopra.

La scoperta di questa regolarità ricorsiva, fatta su un piccolo calcolatore da Feigenbaum, è uno dei principali progressi fatti di recente in questo campo. Essa afferma in particolare che per ottenere la linea $n + 1$ dalla linea n basta lasciar produrre dei «gemelli» a ogni punto della linea n. La nuova generazione dovrebbe avere una densità pari a due volte e mezzo quella della generazione precedente. Detto in termini più corretti, la distanza tra i nuovi gemelli dovrebbe essere α volte più piccola della distanza tra il loro genitore e il suo gemello dove α è una costante approssimativamente uguale a 2,502907875095892 8485... Questa regola diviene sempre più precisa col crescere di n.

E a proposito dei valori delle Λ? Tendono asintoticamente a 1? Abbastanza sorprendente, la risposta è no. Questi valori Λ convergono rapidamente su un particolare valore critico λ_c, pari a circa 0,892486418... e la loro convergenza è notevolmente armoniosa, nel senso che la distanza tra successivi valori Λ si riduce geometricamente. Più precisamente, il rapporto $(\Lambda_n - \Lambda_{n-1})/(\Lambda_{n+1} - \Lambda_n)$ si avvicina a un valore costante chiamato δ da Feigenbaum, il suo scopritore, ma che altri chiamano spesso numero di Feigenbaum. Il suo valore è approssimativamente 4,669201660910299097...

In breve, col tendere di λ a λ_c, per particolari valori di λ ottenuti con la costante δ di Feigenbaum, l'attrattore di f si riproduce e i suoi sempre più numerosi elementi sono disposti sull'asse delle x secondo un semplice schema ricorsivo, il cui principale parametro determinante è l'altra costante di Feigenbaum, α.

Per λ che superano λ_c - quello che viene chiamato regime caotico - l'iterazione di f può, per qualche valore del seme, portare a orbite che non convergono su alcun attrattore finito: si tratta di orbite aperiodiche. Per la maggior parte dei valori del seme, l'orbita rimarrà periodica, ma la sua periodicità sarà molto difficile da individuare. Prima di tutto, il periodo sarà estremamente lungo; in secondo luogo, l'orbita sarà molto più caotica di prima. Una tipica orbita periodica, invece di convergere rapidamente verso un attrattore geometricamente semplice, vagherà sull'intervallo [0,1]; il suo comportamento apparirà indistinguibile dal caos totale e viene indicato col termine ergodico. Inoltre, semi ravvicinati possono dar luogo, con pochissime iterazioni, a orbite del tutto differenti. In breve, una visione statistica dei fenomeni diviene molto più ragionevole se si oltrepassa λ_c.

Ora, cosa hanno a che fare concetti come l'iterazione di funzioni non monotone, il raddoppio del periodo, il regime caotico e così via con lo studio della turbolenza nel flusso idrodinamico, della fluttuazione della popolazione erratica nelle relazioni predatore-preda e dell'instabilità dei modi laser? L'idea base sta nel contrasto tra il flusso laminare e il flusso turbolento. In un fluido che scorre tranquillamente, il flusso è laminare, un termine usato per indicare che tutte le molecole del fluido si muovono come automobili su un'autostrada a più corsie. Le caratteristiche chiave sono (1) che ciascuna automobile segue lo stesso percorso di quella che la precede e (2) che due automobili vicine, sia che si trovino sulla stessa corsia o su corsie differenti, col passar del tempo si separeranno lentamente una dall'altra - essenzialmente in proporzione alle loro velocità, cioè linearmente. Queste caratteristiche si applicano anche alle molecole del fluido in un flusso laminare; in questo caso le corsie sono dette linee di flusso o lamine.

Al contrario, quando un fluido è messo in agitazione da qualche forza esterna, questo tranquillo comportamento diventa turbolento, come per i frangenti sulla spiaggia e la panna mescolata nel caffè. La stessa parola «turbolento» ha un suono molto più ruvido e angoloso di quello della parola «laminare». Qui non vale più l'immagine dell'autostrada a più corsie; le linee di flusso si separano una dall'altra e si aggrovigliano nei modi più contorti. In questi sistemi ci sono gorghi, vortici e spirali di ogni tipo e dimensione contemporaneamente; quindi due punti inizialmente vicinissimi possono trovarsi rapidamente trascinati in regioni del fluido totalmente differenti. Questi percorsi rapidamente divergenti sono il tratto distintivo della turbolenza. La distanza tra punti può avere col tempo un incremento esponenziale, invece che lineare, e il coefficiente del tempo nell'esponente è chiamato numero di Lyapunov. Quando si parla di flusso turbolento, è a questa rapida e quasi imprevedibile separazione dei vicini che ci si riferisce e tale comportamento richiama in modo impressionante la rapida separazione, nel regime caotico di λ, di due orbite anche con semi originariamente vicinissimi.

Questo fatto suggerisce che lo scenario (come viene chiamato), in cui belle orbite periodiche danno luogo alle disordinate orbite caotiche della nostra funzione parabolica, potrebbe forse essere matematicamente identico allo scenario in cui si svolge la transizione alla turbolenza in un fluido o in altri sistemi. Come si possa esattamente stabilire questa connessione, però, richiede qualche precisazione più dettagliata. In particolare, dobbiamo

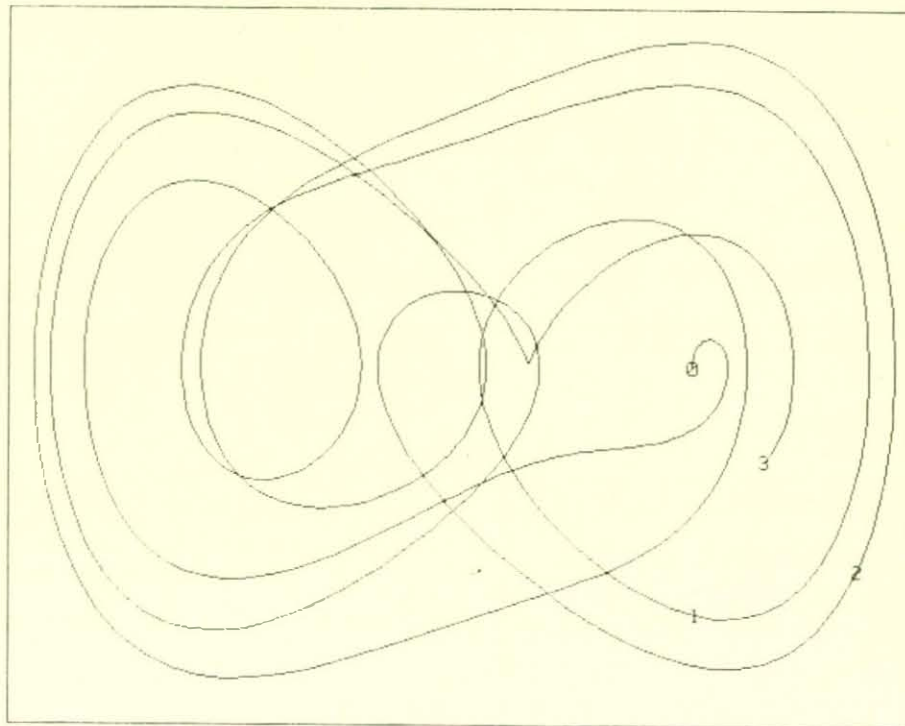
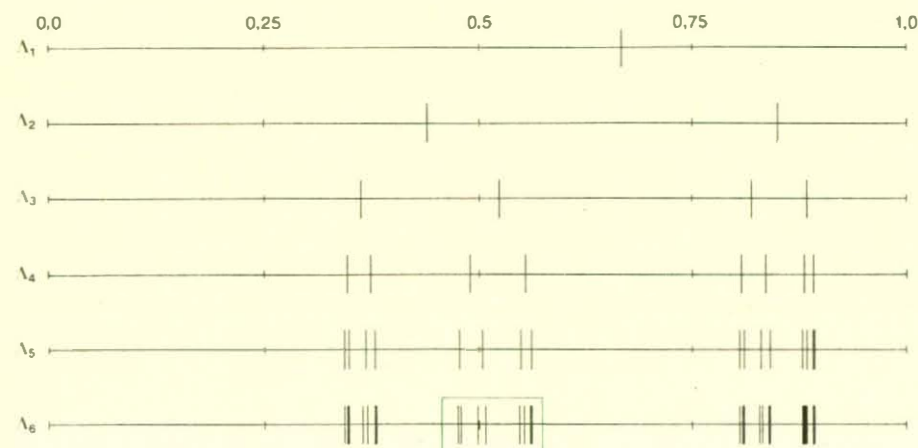
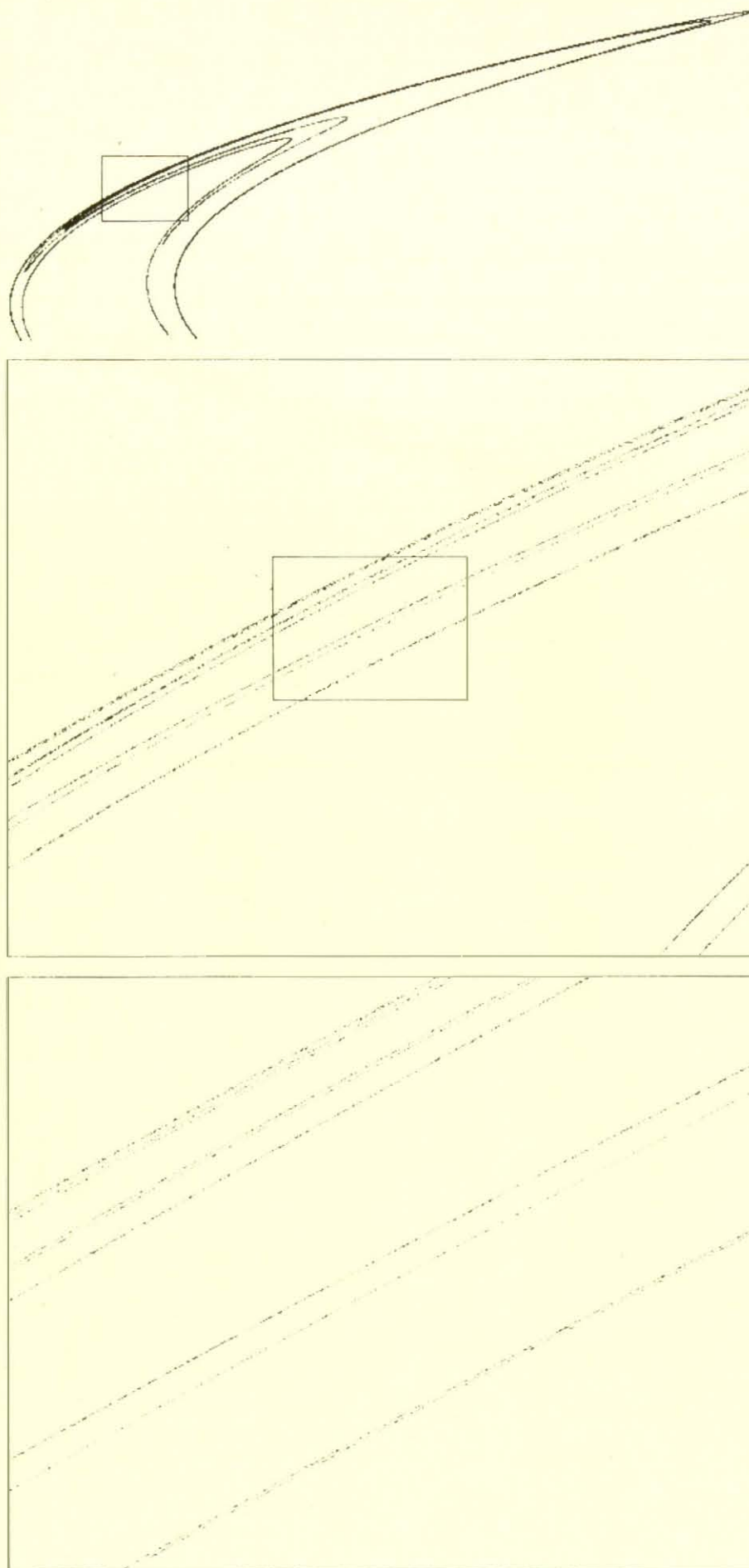
Attrattori stabili divengono instabili e «si scindono» a certi valori di λ **(dall'alto in basso)**

prendere brevemente in considerazione il modo in cui si può tradurre in un modello matematico il flusso spazio-temporale di un fluido o di qualche altra entità, come la densità della popolazione o il denaro.

In questi problemi del mondo reale, le equazioni che finora si sono dimostrate più utili per fornire modelli dei fenomeni sono le equazioni differenziali. Un'equazione differenziale mette in relazione l'entità della variazione di qualche quantità con l'attuale dimensione di quella quantità e con le attuali dimensioni di altre quantità. Inoltre, la stessa variabile tempo è continua; non salta da un istante discreto all'altro come fanno certi strani orologi, ma fluisce in modo indivisibile come un liquido. Un modo per visualizzare le strutture definite dalle equazioni differenziali consiste nell'immaginare uno spazio a più dimensioni - ne abbia pure migliaia o solo alcune - in cui un punto continui a muoversi seguendo una curva. In un qualsiasi momento, il punto contiene tutte le informazioni sullo stato del sistema fisico. Le sue proiezioni lungo i vari assi danno i valori di tutte le quantità rilevanti che definiscono rigorosamente un unico stato. Ovviamente lo spazio - chiamato spazio fase - dovrebbe avere un numero enorme di dimensioni perché un semplice punto possa riassumere la forma di un'onda che si infrange su una spiaggia. D'altra parte, in una semplice relazione predatore-preda sono sufficienti due dimensioni: una variabile, diciamo x, che fornisce la popolazione di predatori e l'altra, diciamo y, che dà la popolazione di prede.

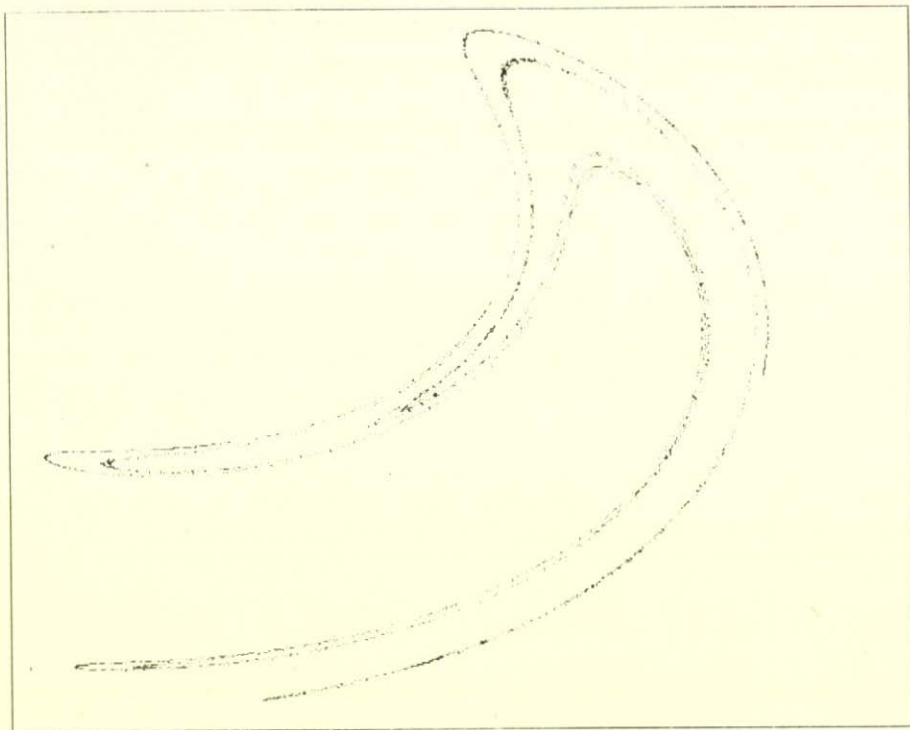
Col progredire del tempo, x e y si determinano l'un l'altro in modo intrecciato. Per esempio, una grande popolazione di predatori tenderà a ridurre la popolazione delle prede, mentre una piccola popolazione di prede tenderà a ridurre la popolazione di predatori. In un sistema di questo tipo, x e y costituiscono un unico punto (x,y) che si avvolge dolcemente in un'orbita continua sul piano. (Qui il senso in cui è usato il termine «orbita» è differente dal precedente, quello dell'orbita discreta, a salti che abbiamo visto quando veniva iterata la nostra funzione parabolica.) Una possibile orbita di questo tipo appare nella figura in alto di questa pagina; essa è generata da un'equazione differenziale chiamata equazione di Duffing. Assomiglia al percorso di una mosca che svolazzi nella vostra camera da letto, o piuttosto, assomiglia all'*ombra* del volo della mosca su una parete. In realtà questa curva intersecantesi a due dimensioni *è* l'ombra di una curva non intersecantesi a tre dimensioni. Il moto di un punto nello spazio fase deve essere *sempre* non intersecantesi; ciò deriva dal fatto che un punto nello spazio fase rappresentante lo stato di un sistema racchiude ogni informazione sul sistema, inclusa la sua storia futura, così che non ci possono essere due differenti percorsi che passino per uno stesso punto.

In particolare, nell'equazione di Duffing c'è una terza variabile, z, che non ho ancora ricordato. Se si pone che x e y rappresentino la popolazione dei predatori e delle prede, allora si può pensare che z rappresenti un'influenza esterna a variazione periodica, come l'azimut del Sole o la quantità di neve sul terreno. Ora, se mi consentite di mescolare la mia immagine della mosca svolazzante con l'esempio dei predatori e delle prede, immaginiamo una mosca che svolazzi periodicamente avanti e indietro tra due pareti. Poniamo che alla mosca occorra un anno per attraversare la stanza e per tornare indietro. (Forse è una stanza da letto piuttosto grande, o forse è un volo decisamente lento.) In ogni caso, mentre la mosca vola la sua ombra su una delle due pareti

La curva generata dall'equazione di Duffing ricorda l'ombra sulla parete del volo di una mosca

I numeri sulla curva segnano le posizioni dell'ombra della mosca a intervalli regolari

Ingrandimento in due stadi dell'attrattore di Hénon che ne rivela dettagli infinitesimali

disegna la curva che si vede nella figura. Se alla mosca capita di tornare in un punto della stanza per cui è già passata, è condannata a girare in tondo per sempre, continuando a seguire il percorso fatto la volta precedente. Questo vi dà un'idea dell'orbita continua di un punto nello spazio fase rappresentante lo stato di un sistema dinamico controllato da equazioni differenziali.

Supponiamo ora di voler stabilire qualche connessione tra questi sistemi e delle orbite *discrete*. Come potremmo fare? Bene, non è necessario tenere d'occhio tutti i momenti i valori x, y e z; essi possono essere controllati periodicamente, secondo qualche frequenza naturale. Nel caso delle popolazioni animali, un anno è l'ovvio periodo naturale. L'azimut del Sole è esattamente periodico e il clima *cerca* almeno di ripetersi di anno in anno. Si può allora scegliere una sequenza naturale di punti discreti (x_1, y_1 e z_1), (x_2, y_2, z_2),... uno per anno. È come se, in una particolare occasione dell'anno, una luce si accendesse e bloccasse la mosca. Per tutto il resto del tempo non si vedono le sue peregrinazioni per la stanza. La figura in basso della pagina precedente mostra una sequenza di punti discreti lungo l'ombra del percorso della mosca, segnata da numeri che ci dicono quando sono apparsi. Gradualmente, passati molti «anni», si sarà accumulato un numero abbastanza grande di questi punti discreti perché inizino essi stessi a dar luogo a una forma riconoscibile. Questo schema di punti è un'«orbita» *discreta* ed è quindi strettamente correlato alle orbite discrete definite dall'iterazione della nostra parabola $f(x)$. Nel caso della parabola avevamo una semplice relazione di ricorrenza unidimensionale (o iterazione): $x_{n+1} = f(x_n)$. Qui abbiamo una ricorrenza bidimensionale: $x_{n+1} = f_1(x_n, y_n)$ e $y_{n+1} = f_2(x_n, y_n)$. Questo è un sistema di relazioni di ricorrenza accoppiate, in cui i valori di output dell'n-esima generazione (x_n, y_n) sono reintrodotti come nuovi input in f_1 e f_2, per produrre la $n + 1$-esima generazione; e la cosa va sempre avanti, generazione dopo generazione. Naturalmente, se aumentano le dimensioni c'è un numero molto maggiore di tali equazioni. Nondimeno, lo schema di tutti questi sistemi rimane lo stesso: un punto multidimensionale (x_n, y_n, z_n...) salta da una posizione discreta a un'altra nello spazio fase, man mano che viene accresciuta una variabile discreta (n, che rappresenta il tempo che procede a salti in unità discrete).

Si noti che noi abbiamo semplificato la nostra osservazione della variabile continua tempo che è implicata nelle equazioni differenziali. Lo abbiamo fatto focalizzando la nostra attenzione sul modo in cui il punto si connette al suo predecessore dell'«anno» prima (o qualunque sia il periodo naturale considerato). Ma c'è sempre un «periodo naturale» da osservare in un sistema di equazioni differenziali mutuamente intrecciate? Non sempre. In alcune situazioni però c'è, ed è questo il caso di tutte le situazioni in cui si ha comportamento turbolento.

L'equazione di Duffing genera uno strano attrattore

Perché ciò avviene? Tutti i sistemi che mostrano comportamento turbolento sono «dispersivi», il che significa che essi disperdono, o degradano, energia da forme più utilizzabili, quale l'elettricità, nella meno utilizzabile forma del calore. Ciò è causato dall'attrito, nel caso del flusso idrodinamico, e da analoghi derivati dell'attrito negli altri sistemi che abbiamo preso in considerazione. Una familiare conseguenza dell'attrito è che il moto degli oggetti si arresta a meno che non venga immessa energia. Ora, se noi «mettiamo in azione» un sistema dispersivo con una forza *periodica* (si può immaginare, per esempio, di mescolare una tazzina di caffè con un cucchiaino in modo periodico e circolare), allora il sistema ovviamente non si arresterà e si muoverà verso un qualche tipo di stato regolare. Tale stato regolare è un'orbita stabile, o, per usare una parola che ci è familiare, un attrattore nello spazio fase. Dato, però, che abbiamo messo in moto il sistema con un cucchiaino periodico, abbiamo definito una frequenza naturale alla quale far lampeggiare la nostra luce e bloccare lo stato del sistema, vale a dire ogni volta che il cucchiaino arriva a passare nella sua rotazione per qualche dato segno sulla tazzina, ad esempio il manico. Questo costituirà il nostro «anno». In questo modo il tempo continuo può essere sostituito da una serie di istanti discreti, finché abbiamo a che fare con un sistema dispersivo guidato da una forza periodica. E così si possono rimpiazzare orbite continue con orbite discrete, ciò che riporta nel quadro l'iterazione.

Se la forza di guida non ha alcun periodo naturale (può essere semplicemente una forza costante), c'è ancora un modo per definire un periodo naturale, nel caso che qualche variabile del sistema oscilli avanti e indietro tra degli estremi: basta far scattare la luce ogni volta che la variabile tocca il suo valore estremo, e la mosca sarà ancora colta in istanti discreti. Questo tipo di rappresentazione discreta del moto di una mosca in uno spazio a più dimensioni è chiamato comunemente mappa di Poincaré.

L'esempio della tazzina non è naturalmente abbastanza rigoroso per convincere un matematico; eppure dà un'idea di come lo studio di un insieme di equazioni differenziali accoppiate possa essere rimpiazzato dallo studio di un insieme di relazioni discrete di ricorrenza accoppiate. Questo è il passo fondamentale che ci riporta alle recenti scoperte a proposito della parabola.

Nel 1975, Feigenbaum scoprì che i suoi numeri α e δ non dipendono in realtà dai dettagli della forma della curva definita da $f(x)$. La cosa sarebbe uguale quasi per tutte le forme dolcemente convesse con picco nello stesso punto. Ispirato dall'universalità strutturale scoperta da Metropolis, Stein e Stein, Feigenbaum cercò di lavorare con una curva sinusoidale invece che con una parabola e rimase sbalordito nel vedere riapparire gli stessi valori numerici, uguali per molti decimali, dei numeri α e δ che avevano caratterizzato il raddoppio del periodo e l'inizio del caos per la parabola. Proprio come per la parabola, c'è per la curva sinusoidale un parametro d'altezza λ e un insieme di speciali valori λ che convergono in un punto critico λ_c. Inoltre, l'inizio del caos in λ_c è governato dagli stessi numeri α e δ. Feigenbaum iniziò a sospettare che ci fosse qualcosa di universale. In altre parole, sospettò che più importante di f stessa è il fatto che f continui a essere iterata e che anzi f non abbia alcun ruolo nell'inizio del caos.

La cosa, naturalmente, non è così semplice. Feigenbaum scoprì presto che l'importante di f è la natura del picco, proprio al suo centro. Sul lungo periodo, il comportamento delle orbite dipende solo da un frammento infinitesimale della cresta del grafo e infine dipende solo dal comportamento proprio nel punto in cui si raggiunge il massimo. Il resto della forma, anche le regioni più prossime al vertice, è irrilevante. Una parabola ha ciò che viene chiamato un massimo quadratico, e lo stesso vale per un'onda sinusoidale, un cerchio e un ellisse. In effetti, ci si dovrebbe aspettare che il comportamento di una funzione a cui corrisponde una curvatura dolce e prodotta in modo casuale, in prossimità di un massimo tipico, sia di tipo quadratico, in assenza di particolari coincidenze. Quindi il caso della parabola, piuttosto che una bizzarra eccezione inizia ad apparire la regola. Questa scoperta empirica di Feigenbaum, in cui sono implicati due fondamentali fattori di graduazione, α e δ, che caratterizzano l'inizio del caos attraverso attrattori a raddoppio di periodo, rappresenta un nuovo tipo di universalità, nota come universalità *metrica* per distinguerla dalla già conosciuta universalità strutturale. Oscar E. Lanford III ha dimostrato (nel senso più tradizionale della dimostrazione) la correttezza di questa universalità metrica per il caso unidimensionale.

Uno sviluppo veramente entusiasmante si ebbe quando le costanti di Feigenbaum si presentarono inaspettatamente in alcuni modelli caotici di effettivi sistemi fisici, non in bei sistemi matematici idealizzati. Walter Franceschini, dell'Università di Modena, adattò per la simulazione col calcolatore l'equazione di Navier-Stokes, che regola tutti i flussi idrodinamici. Per farlo, trasformò l'equazione in un insieme di cinque equazioni differenziali accoppiate di cui poteva poi studiare numericamente sul calcolatore le mappe di Poincaré. Egli scoprì per prima cosa che il sistema mostrava attrattori con ripetuti raddoppi del periodo via via che i suoi parametri regolativi si avvicinavano ai valori in cui ci si aspettava la turbolenza. All'oscuro del lavoro di Feigenbaum, egli mostrò i suoi risultati a Jean-Pierre Eckmann, dell'Università di Ginevra, che immediatamente lo spinse a stabilire il tasso di convergenza dei valori λ secondo cui avveniva il raddoppiamento del periodo. Con loro grande sorpresa, i valori α e δ di Feigenbaum - precisi a circa quattro decimali - saltarono fuori apparentemente dal nulla! Per la prima volta, un accurato modello matematico di reale turbolenza fisica rivelava che la sua struttura era intimamente connessa al modesto caos nascosto nella modesta parabola $y = 4\lambda x(1-x)$. Successivamente, Eckmann, Pierre Collet e H. Koch dimostrarono che in un sistema dispersivo multidimensionale guidato, tutte le dimensioni tranne una tendono a scomparire dopo un periodo di tempo sufficientemente lungo e quindi ci si dovrebbe *aspettare* che riappaiano le caratteristiche del comportamento unidimensionale, vale a dire l'universalità metrica di Feigenbaum.

Da allora, gli sperimentatori hanno tenuto gli occhi ben aperti per individuare comportamenti di raddoppio del periodo in reali sistemi fisici (non in modelli prodotti dal calcolatore). Un comportamento di questo tipo è stato osservato in certi tipi di flusso convettivo, ma finora le misure sono troppo imprecise per fornire un valido sostegno all'idea che la parabola contenga le indicazioni necessarie per individuare la reale natura della turbolenza fisica. Certo è allettante pensare che in qualche modo tutto ciò che importa è che un insieme dispersivo di relazioni di ricorrenza accoppiate venga iterato, mentre i dettagli delle proprietà di quelle ricorrenze possono essere totalmente ignorati se ci si concentra sul capire come si giunge alla turbolenza.

Feigenbaum mette la cosa in questi termini. Spesso si vede in cielo una configurazione di nuvole - un reticolato celeste composto da una miriade di sbuffi bianchi stesi da orizzonte a orizzonte - che chiaramente non si è presentata «per caso»: deve avere operato qualche legge idrodinamica sistematica. Eppure, dice Feigenbaum, deve esserci una legge che opera a un livello superiore, ossia su scala maggiore, dell'equazione di Navier-Stokes, che si basa su volumi infinitesimali di fluido e non su grossi «blocchi». Sembra che per capire quelle belle configurazioni in cielo si debba in qualche modo superare i *dettagli* dell'equazione di Navier-Stokes e arrivare a qualche modo più grossolano ma anche più rilevante di analizzare il flusso idrodinamico. La scoperta che l'iterazione dà luogo a universalità - cioè indipendenza dai dettagli della funzione (o funzioni) iterata - fa sperare che questa concezione dell'idrodinamica stia per emergere.

Abbiamo parlato di attrattori e di turbolenza; e gli strani attrattori? Ormai abbiamo dato i concetti necessari per comprendere questa nozione. Quando un sistema dispersivo a due dimensioni (o a più dimensioni) periodicamente guidato ha come modello un insieme di iterazioni accoppiate, l'insieme dei punti catturati dai successivi lampi della luce periodica disegna una forma che gioca per questo sistema il ruolo che un'orbita semplice aveva per la nostra parabola. Ora, le possibilità sono maggiori quando si opera in uno spazio con più di una dimensione. Certo è possibile avere un punto fisso stabile, o un attrattore di periodo uno. Questo significherebbe che a ogni lampo di luce il punto che rappresenta lo stato del sistema è esattamente dov'era l'ultima volta. È anche possibile avere un attrattore periodico: uno in cui dopo un numero finito di lampi il punto è ritornato su una posizione precedente. La cosa sarebbe analoga ai cicli-2, cicli-4 e così via che abbiamo visto presentarsi nel caso della parabola.

C'è tuttavia un'altra possibilità: quella che il punto non ritorni mai nella sua posizione originaria nello spazio fase; i successivi lampi di luce mostrano che esso compie un cammino erratico all'interno di una ristretta regione dello spazio fase. Dopo un certo periodo di tempo questa regione, illuminata periodicamente dalla luce, può prendere forma di fronte agli occhi dell'osservatore. Nella maggioranza dei casi finora osservati ha luogo un fenomeno del tutto inaspettato: il punto che saltella in modo casuale crea in realtà una delicata filigrana che ricorda «le fantastiche e appena percettibili tracce che il gelo lascia sul vetro». (Devo questa immagine poetica al critico americano James Huneker che se ne è servito per descrivere il magico effetto prodotto da uno degli studi per pianoforte di Chopin, il n. 2, Op. 25.) La delicatezza di tali forme è di un tipo abbastanza preciso e ricorda da vicino le curve «fratte» descritte da Benoit Mandelbrot nel suo libro *Fractals: Form, Chance, and Dimension*. In particolare, qualsiasi sezione di un attrattore cosiffatto si rivela altrettanto squisitamente precisa nei dettagli quanto il più ampio quadro da cui è stata estratta. In altre parole c'è un infinito ritorno di dettagli, un intreccio senza fine di configurazioni nelle configurazioni. Nella figura di pagina 102 si può vedere una delle prime di tali strutture, detta l'attrattore di Hénon. È generata dalla serie di punti (x_n, y_n) definiti dalle relazioni di ricorrenza $x_{n+1} = y_n - ax_n^2 - 1$ e $y_{n+1} = bx_n$. Qui a è uguale a 7/5 e b a 3/10; i valori del seme sono $x_0 = 0$ e $y_0 = 0$. Il piccolo quadrato della figura in alto è ingrandito nella figura di mezzo per mettere meglio in luce i dettagli e un altro quadrato nella figura di mezzo è pure ingrandito per rivelare dettagli di dimensioni ancora minori. Si noti che quello che risulta assomiglia a un'autostrada a tre corsie ciascuna delle quali, quando viene ingrandita, si divide in più corsie parallele la più estrema delle quali è una nuova autostrada a tre corsie e così via. Qualsiasi sezione incrociata perpendicolare di tale autostrada darebbe un insieme di Cantor, formato con un semplice e ben noto procedimento ricorsivo.

Si cominci con un intervallo chiuso, diciamo [0,1]. («Chiuso» significa che l'intervallo comprende i suoi estremi.) Si elimini ora un sotto-intervallo aperto centrale. (Dato che un sotto-intervallo aperto non comprende i suoi estremi, quei due punti rimarranno nell'insieme cantoriano, costruiti sotto i vostri occhi.) Di solito si sceglie come sotto-intervallo il terzo di mezzo (1/3, 2/3), ma non è necessario. Rimangono due sotto-intervalli chiusi. Li si sottoponga allo stesso tipo di procedimento, eliminando al loro interno un sotto-intervallo aperto centrale. Si ripeta il procedimento all'infinito: al termine di questa impresa infinita rimarrà una delicata struttura formata da punti isolati disposti lungo il segmento originale [0,1] come gocce di rugiada su un filo. Essi, però saranno più che numerabili e la loro densità dipenderà dai dettagli del processo ricorsivo di eliminazione. È questa la natura di un insieme di Cantor, e se le sezioni incrociate di un attrattore hanno questa bizzarra distribuzione, l'attrattore è detto (e a ragione) strano.

Un altro interessante strano attrattore è generato dai punti «stroboscopici» 0, 1, 2,... della figura a pagina 101 in basso. Dato che questa configurazione è ricavata dall'equazione di Duffing, l'attrattore è detto di Duffing e lo si può vedere in scala leggermente maggiore nella figura della pagina a fronte. Si noti la sua marcata somiglianza con l'attrattore di Hénon; forse anche qui si riaffaccia l'universalità.

È interessante il fatto che per la parabola, al valore critico λ_c, l'attrattore di f diviene aperiodico ed è formato da un numero infinito di punti. (Dopo tutto, è il culmine di una serie infinita di raddoppi di periodo.) Inoltre, la disposizione di quegli infiniti punti sull'intervallo [0,1] è stata costruita con la regola ricorsiva di Feigenbaum che, come si è visto nella figura di pagina 100, comporta la costante α. Non è difficile credere che ciò implichi che questo particolare attrattore sia esso stesso un insieme di Cantor. La fertile parabola ci ha quindi fornito un esempio di strano attrattore *unidimensionale*!

Nel regime caotico del più generale caso k-dimensionale, è del tutto impossibile una predizione a lungo termine del percorso di un punto: due punti che quasi si tocchino su uno strano attrattore, dopo pochi lampi di luce sono finiti in posti totalmente diversi. Questo fatto è detto «dipendenza sensibile dalle condizioni iniziali» ed è un altro criterio definitorio di uno strano attrattore.

Allo stato attuale delle conoscenze, nessuno sa di preciso perché, come e quando saltino fuori strani attrattori nei regimi caotici che rappresentano sistemi fisici dispersivi, ma certo essi sembrano svolgere un ruolo centrale nel mistero della turbolenza. David Ruelle, uno dei primi a seguire questo approccio al problema della turbolenza, ha scritto («Questi sistemi di curve, queste nuvole di punti, a volte evocano galassie di fuochi artificiali, altre volte bizzarre e fastidiose inflorescenze. C'è un intero mondo di forme ancora da esplorare e di armonie ancora da scoprire.»

Robert M. May, uno studioso di biologia teorica, in un ormai famoso articolo panoramico sull'argomento apparso nel 1976, concludeva con una perorazione che trovo molto appropriata e che vorrei riportare:

«Vorrei... che a chi studia matematica venisse presto sottoposta l'equazione $y = 4\lambda x(1-x)$. Questa equazione può essere studiata da un punto di vista fenomenologico iterandola su un calcolatore o anche manualmente. Il suo studio non comporta la sofisticazione concettuale del calcolo elementare e sarebbe di grande aiuto allo studente per capire i sistemi non lineari.

«Non solo nel campo della ricerca, ma anche nel mondo quotidiano della politica e dell'economia staremmo tutti di gran lunga meglio se più persone capissero che semplici sistemi non lineari non possiedono necessariamente semplici proprietà dinamiche».

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Alla scoperta delle strane attrattive del caos

Il caos ha strane attrattive per la mente che riesca a vedervi degli schemi. Così avviene per alcuni sistemi fisici che mostrano un comportamento caotico e sono in un certo senso attirati verso questi schemi. Per di più, gli schemi stessi hanno una strana attrattiva. Forse alcuni lettori già sanno che le forme geometriche sottostanti al caos si chiamano attrattori strani, o caotici (si veda in proposito l'articolo *Il caos* di James P. Crutchfield, J. Doyne Farmer, Norman H. Packard e Robert S. Shaw in «Le Scienze» n. 222, febbraio 1987). Si possono generare attrattori strani con un calcolatore domestico.

Prima di iniziare a seguirmi, i lettori si devono munire di un rivestimento protettivo di intuizione fisica. In particolare, che cos'è un attrattore? Grosso modo, un attrattore è una generalizzazione della nozione di equilibrio; è ciò su cui si assesta, o verso cui è attratto, il comportamento di un sistema. Il pendolo è un semplice sistema fisico che illustra il concetto di attrattore. Prendiamo un comune pendolo che si muove sottostando a forze di attrito che lo rallentano fino a farlo fermare. Si può descrivere il movimento del pendolo utilizzando un cosiddetto diagramma di fase, o di stato, in cui vengono rappresentati su un asse l'angolo del pendolo con la verticale e sull'altro la velocità con cui cambia l'angolo. Il moto oscillante del pendolo è rappresentato da un punto che gira intorno all'origine del diagramma di fase; man mano che il pendolo perde energia, il punto si avvicina a spirale all'origine, dove infine va a fermarsi. In questo caso l'origine viene chiamata un attrattore perché sembra attrarre il punto in movimento nel diagramma di fase. Giustamente i lettori penseranno che non ci sia niente di strano in un attrattore formato da un singolo punto.

Leggermente più complesso è l'attrattore insito nel movimento di una pendola a muro, in cui un meccanismo di scappamento fornisce energia al pendolo impedendogli di rallentare. Se si fa partire l'orologio con una spinta troppo energica al pendolo, questo rallenta fino al ritmo prescritto dallo scappamento ma poi non rallenta più. Se invece si avvia l'orologio con una spinta troppo debole, il pendolo si comporta nel solito modo: rallenta fino a fermarsi. Nel caso della spinta troppo energica, il moto del pendolo in un diagramma di fase è una spirale che si muove a spire sempre più strette intorno a un'orbita circolare. L'attrattore qui è una curva circolare e in questo contesto un cerchio non è più strano di un punto.

Si può far assumere a un comune pendolo un comportamento caotico introducendo un movimento di vibrazione verticale: se con un motore elettrico si sposta su e giù in modo sinusoidale il punto di sostegno, è possibile che il pendolo inizi a oscillare in modo scomposto senza mostrare più alcun comportamento periodico.

Per parlare del caos, però, ho scelto un diverso sistema fisico. Immaginiamo di avere tre amplificatori, il primo dei quali dà in uscita un segnale x che viene passato agli altri due; il secondo amplificatore dà in uscita il segnale $1 - x$ in risposta a x. Il terzo amplificatore assume in ingresso i due segnali, x e $1 - x$, genera il loro prodotto $x(1 - x)$ e lo passa al primo amplificatore, il quale riceve in ingresso anche una tensione di controllo r. Il circuito è completato da un componente aggiuntivo, inserito sulla linea di uscita del primo amplificatore, che campiona il proprio ingresso ed emette per breve tempo la stessa tensione. Il circuito formato dai tre amplificatori esegue una «danza di tensione» che diventa sempre più sfrenata con l'aumentare della tensione di controllo r.

Quando r è minore di 3 e x ha all'inizio un valore diverso da zero, il circuito ha una breve oscillazione prima di assestarsi su un certo valore di x che rimane poi immutato. Questo valore costituisce un attrattore a un unico punto. Se ora si alza la tensione di controllo r a un livello subito superiore a 3, il circuito oscilla tra due valori di x. A questo livello di r, il circuito è detto bistabile e l'attrattore è formato da due punti. Se si aumenta ancora r, il circuito oscilla tra quattro punti; un ulteriore aumento produce un attrattore a otto punti. Portando su valori sempre più alti la manopola che controlla r, continua a riprodursi lo schema fatto di raddoppi, finché improvvisamente, in una posizione grosso modo a metà strada tra 3 e 4, il circuito impazzisce e parte per una ricerca senza fine, a velocità elettronica, del semplice schema ricorrente che caratterizzava la sua esistenza precedente. Il suo comportamento ora è governato da un attrattore strano con valori potenzialmente infiniti. Il risultato è il caos.

Forse i lettori esperti di elettronica sono tentati di costruire un circuito del genere. Altri possono simularlo su un calcolatore di qualsiasi dimensione e osservare la danza riprodursi con grande chiarezza sullo schermo. Per ottenere questo risultato devono scrivere un semplice programma che calcoli l'equazione iterata $x \leftarrow rx(1 - x)$. Il programma, che io chiamo CHAOS1, ha un nucleo formato da sei istruzioni:

```
x ← 0,3
for i ← 1 to 200
    x ← rx(1 − x)
for i ← 1 to 300
    x ← rx(1 − x)
    tracciare (200x, 100)
```

Il valore iniziale della variabile x è 0,3. CHAOS1 entra poi in un ciclo che ripete 200 volte l'equazione base per far sparire i transitori. I transitori sono inerenti all'equazione stessa, non a un'imprecisione aritmetica. Le ragioni di questo fenomeno saranno chiarite più avanti in termini geometrici. Il programma entra poi in un nuovo ciclo che ripete altre 300 volte l'equazione, tracciando ogni volta il valore di x.

Il numero 100 usato nella precedente istruzione di tracciamento ha un valore più generico che specifico; qui lo schermo ha una dimensione di 200 per 200. La coordinata orizzontale, $200x$, distribuisce i diversi valori calcolati per x (sempre compresi tra 0 e 1) lungo una riga dello schermo, posta a un'altezza 100, cioè a metà dello schermo ipotetico.

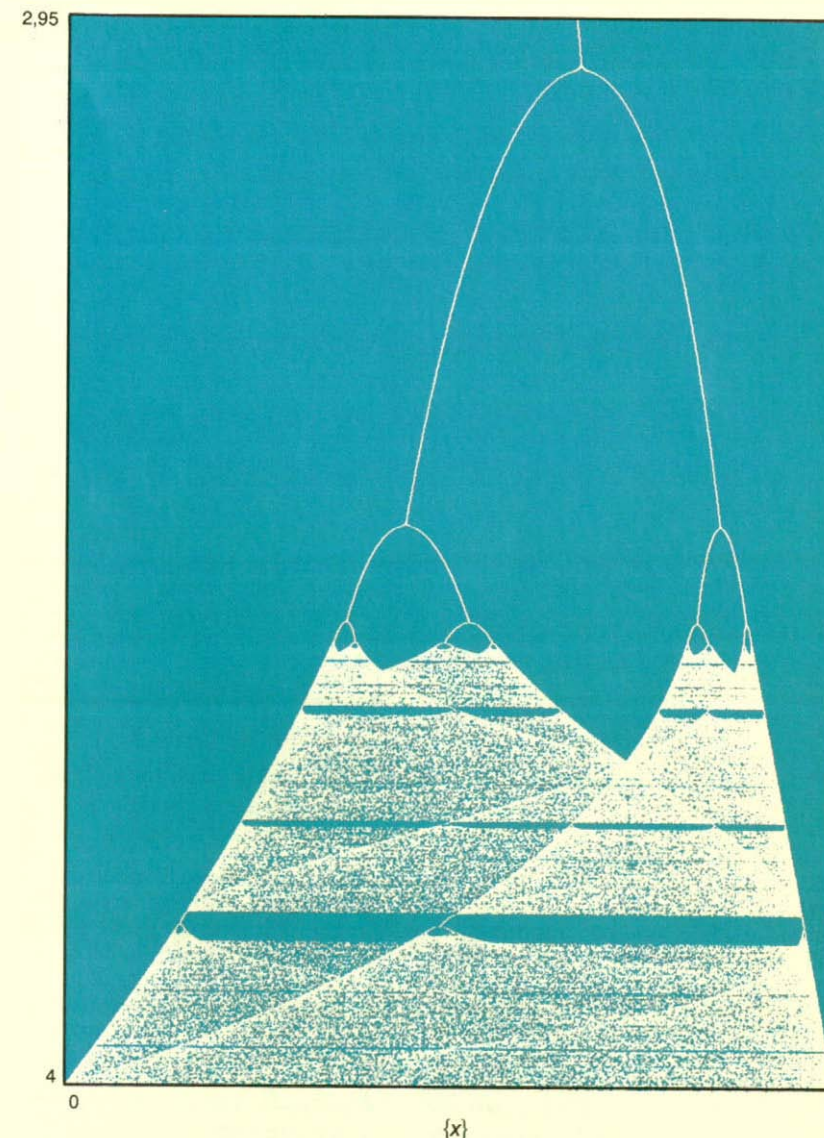
A seconda del valore della variabile di controllo r, il nucleo del programma traccerà un unico punto 300 volte oppure diversi punti meno di 300 volte ciascuno. Può anche cercare di catturare il caos tracciando 300 diversi punti di un attrattore strano, che può essere reso più visibile alzando il limite di iterazione. In tutti i casi, una volta stabilizzatosi il processo di iterazione, i valori di x saltano in modo sistematico da un punto dell'attrattore a un altro. Gli attrattori sono anche chiamati orbite, indipendentemente dal fatto che abbiano un numero finito o infinito di punti.

Si può avere una rappresentazione completa del comportamento del semplice circuito di amplificatori se il programma calcola una massa di punti, ciascuno tracciato sotto il precedente (*si veda l'illustrazione nella pagina a fronte*). I punti risultano da una successione di valori di r che vanno da 2,9 a 4,0 in, diciamo, 200 passi a partire dall'estremità superiore dello schermo per arrivare a quella inferiore. Una rappresentazione più elegante si ha usando un numero maggiore di passi, per esempio 4000, ma in questo caso il diagramma non starà nello schermo e deve essere stampato per poter essere visto per intero.

Per valori di r inferiori a 3,56 (più precisamente il valore è 3,56994571869) gli attrattori del semplice sistema dinamico rappresentato nell'equazione iterata $x \leftarrow rx(1 - x)$ sono formati da pochi punti. Questi punti, che rappresentano un comportamento non caotico, sono disposti in tre larghe bande e in un'infinità di bande più piccole. Gli attrattori diventano strani con l'approssimarsi di r a 3,56. Il caos si instaura quando le linee, che fino a quel punto si biforcavano in modo regolare, diventano improvvisamente una folle alternanza di punti bianchi e neri. Abbastanza stranamente, di tanto in tanto il regime caotico scompare mentre r continua la sua inesorabile marcia verso il 4.

Il tracciato completo è chiamato diagramma di biforcazione. Se osservato lateralmente, assomiglia allo spettro del caos visto da una stella di nome x. Il grafico è abbellito da curve e da pieghe con attraenti ombreggiature. Le ragioni dei dettagli ornamentali sono misteri che possono essere spiegati solo dalla teoria del caos. Approfondirò questo argomento più avanti. Per ora c'è un mistero più vicino alla mente di molti lettori: perché quell'equazione dall'aspetto così innocente si comporta in modo così strano?

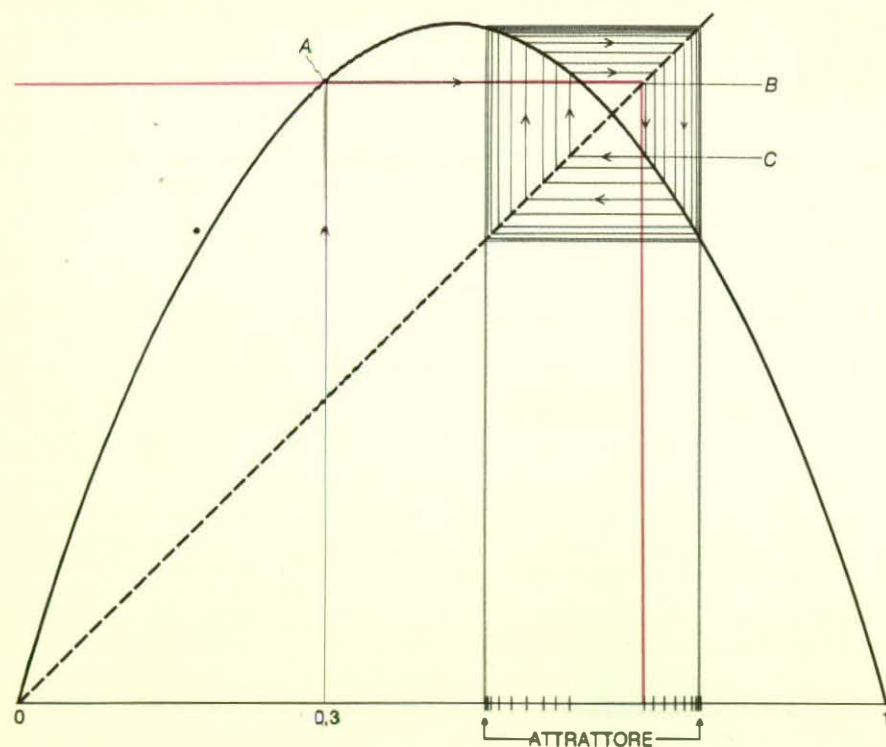
Il comportamento dell'equazione per valori non caotici di r può essere simulato geometricamente disegnando una parabola rappresentata dall'equazione $y = rx(1 - x)$, dove x è la variabile orizzontale e y quella verticale. Sovrapponiamo ora alla parabola la diagonale $y = x$. Un procedimento del genere è stato seguito nell'illustrazione in alto nella pagina successiva, dove r è posto uguale a 3,3, un valore per il quale l'attrattore del sistema è formato da due punti. Per mostrare come si comporta il sistema, si sceglie un valore iniziale di x. Io ho scelto 0,3, ma quasi qualsiasi altro valore sarebbe andato bene.

La prima iterazione dell'equazione è simulata disegnando una linea verticale che inizia al punto $x = 0,3$ nella parte inferiore del grafico e proseguendola finché raggiunge la parabola. Ho identificato con A il punto in cui la retta tocca la parabola. L'altezza dell'intersezione determina il corrispondente valore di y. Nella seconda iterazione quel valore di y viene restituito all'equazione come variabile x. Da un punto di vista grafico, il procedimento equivale a misurare l'altezza dell'intersezione, identificarla con un segno sull'asse orizzontale e disegnare un'altra linea verticale che da quel segno va a raggiungere la parabola. A questo punto si segue una scorciatoia disegnando una linea orizzontale che va dal punto A alla diagonale $y = x$; chiamo B il nuovo punto d'intersezione. Si noti

Un diagramma di biforcazione mostra la transizione al caos

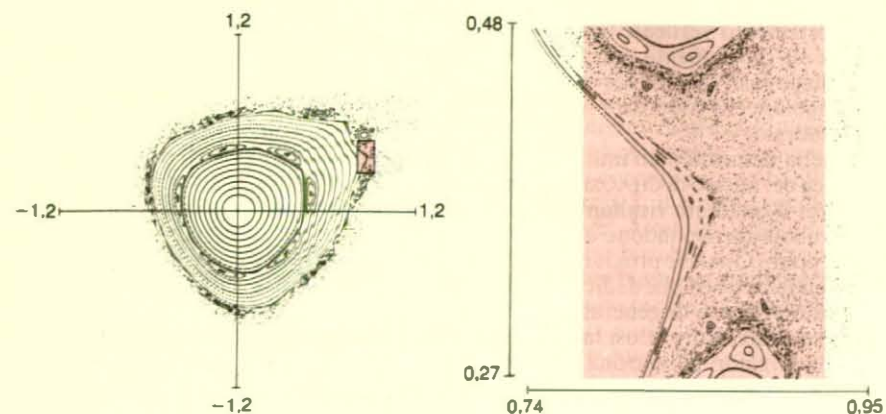
che il punto B e l'origine si trovano in vertici diagonalmente opposti di un quadrato con lati di lunghezza pari al valore di y determinato con la prima iterazione. Di conseguenza, si può reinserire il valore y nel sistema disegnando una linea verticale da B alla parabola (punto C). Continuando a ripetere il procedimento di spostamento verticale fino alla parabola e di spostamento orizzontale fino alla diagonale, si produce un percorso rettangolare che procede a spirale verso un quadrato.

Quella descritta è un'imitazione geometrica del nucleo di CHAOS1. I due punti in cui il quadrato risultante interseca la parabola corrispondono all'attrattore a due punti. Qualche programmatore intraprendente potrebbe dedicarsi all'interessante progetto di generare queste figure col calcolatore. Così facendo i nostri investigatori in poltrona potrebbero scoprire qualche cosa di nuovo sulla «semplice» equazione iterativa in esame. In particolare, che aspetto hanno le figure quando si instaura il caos? Sono i numeri apparentemente casuali generati dai valori di r che producono un caos realmente casuale?

Devo l'idea di un'escursione nel caos a un gran numero di lettori che mi hanno scritto al proposito. Tra loro c'era James P. Crutchfield, uno degli autori dell'articolo *Il caos* precedentemente citato. Crutchfield e gli altri autori spiegano che «la chiave per interpretare il comportamento caotico sta nella comprensione di una semplice operazione di stiramento e piegatura, che ha luogo nello spazio degli stati». Nel caso del semplice sistema di amplificatori, lo spazio degli stati è un segmento che contiene i punti dell'attrattore e il punto che rappresenta il valore corrente di x. Dove entrano in gioco lo stiramento e la piegatura? (Si veda in particolare, sempre del succitato articolo, la pagina 15)

Iterare l'equazione $x \leftarrow rx(1 - x)$ equivale a rappresentare i punti tra 0 e 1 su una curva parabolica. I punti vicini

In una simulazione geometrica di un semplice sistema appare un attrattore a due punti

uno all'altro nell'intervallo unità, soprattutto quelli vicini a 0, finiscono più lontani quando vengono rappresentati sulla curva parabolica. Questa situazione si verifica, naturalmente, quando il numero $rx(1 - x)$ sostituisce x. L'operazione di piegatura interviene a causa della simmetria bilaterale della parabola; se si eccettua l'apice della curva, ci sono sempre due punti dell'intervallo unità che rappresentano lo stesso valore $rx(1 - x)$. Questi punti sono ovviamente x e $1 - x$.

La struttura del diagramma di biforcazione è stata in buona parte analizzata dagli studiosi del caos. I confini delle regioni caotiche sono stabiliti dai valori minimo e massimo delle iterazioni di $x = 0{,}5$. Le curve seguite dai minimi e dai massimi, così come quelle seguite dai «veli» che scendono in modo così strano nelle regioni caotiche, sono tutte dei semplici polinomi in r. Nelle posizioni in cui l'ombreggiatura è più densa si trova la concentrazione più alta di punti negli attrattori strani che l'attraversano. Nelle bande vuote a cui avevo accennato in precedenza, il caos lascia il passo all'ordine. La teoria ci dice che per ogni numero intero c'è una banda (sia pure stretta) con orbite che hanno esattamente quella dimensione. Infine non risulterà sorprendente, a quei lettori che abbiano familiarità con il caos, il fatto che gli attrattori strani, anche nell'umile sistema che abbiamo appena indagato, abbiano una natura frattale; un numero infinito di punti mostra interessanti dettagli a tutti i livelli di ingrandimento, come l'insieme di Mandelbrot che è stato descritto in questa rubrica nell'ottobre 1985.

Orbite successive di una rappresentazione di Hénon **(a sinistra)** *degenerano nel caos* **(a destra)**

Sistemi dinamici più complessi sono incorporati nelle equazioni che prendono il nome da Michel Hénon, un matematico francese. Le cosiddette correlazioni di Hénon non solo descrivono sistemi fisici quali asteroidi in movimento e rubinetti sgocciolanti, ma generano anche, nel corso della loro elaborazione, belle immagini. Una correlazione di Hénon è formata non da un'equazione ma da due. Ecco un esempio:

$$x \leftarrow x\cos(a) - (y - x^2)\mathrm{sen}(a)$$
$$y \leftarrow x\mathrm{sen}(a) + (y - x^2)\cos(a)$$

I valori correnti di due variabili, x e y, vengono usati nella parte destra delle due equazioni per produrre nuovi valori (anch'essi simbolizzati con x e y) nella parte sinistra.

Un secondo programma, chiamato CHAOS2, sfrutta le due equazioni per produrre immagini dell'ordine e del caos insiti in un'ampia classe di sistemi dinamici. CHAOS2 ha un nucleo di programma analogo a quello di CHAOS1:

```
input x e y
for i ← 1 to 1000
    xx ← xcos(a) − (y − x²)sen(a)
    y ← xsen(a) + (y − x²)cos(a)
    x ← xx
    tracciare (100x, 100y)
```

Le ragioni delle differenze tra i due nuclei di programma sono di due tipi: CHAOS2 ha due variabili ripetute invece di una e il sistema descritto dalla correlazione di Hénon è conservativo anziché dissipativo. La presenza di due variabili obbliga a utilizzare una variabile temporanea xx per indicare il nuovo valore di x mentre nella seconda equazione si usa ancora il suo valore corrente x. Il fatto che il sottostante sistema dinamico sia conservativo significa che si può togliere il ciclo di iterazione principale per eliminare valori transitori. Non ci sono perdite di energia dovute ad attrito o altre dispersioni. Di conseguenza non ci sono attrattori veri e propri. Si potrebbe dire, però, che ogni orbita calcolata dal sistema sia il proprio stesso attrattore. In ogni caso, la stranezza (e il caos) è certamente presente nella correlazione di Hénon. Infine, per ogni valore del parametro a il sistema risultante ha una grande quantità di orbite e, a causa della tendenza conservativa, qualsiasi coppia iniziale di valori per x e y rappresenterà un punto che si trova già su una delle orbite; l'attrazione è immediata, per così dire. Per questo motivo, il nucleo di CHAOS2 non usa valori iniziali standard per le sue variabili di iterazione: questi valori devono essere forniti alla macchina dal programmatore.

CHAOS2 è completo quando il suo nucleo è preceduto da un'istruzione di ingresso che permette al programmatore di scegliere il valore di a. Come in

La rappresentazione di Hénon genera figure diverse per a = *0,264* (a sinistra) *e* a = *1,5732* (a destra)

CHAOS1, ciascun nuovo valore di a porta a un nuovo sistema. Ma dato che il sistema è bidimensionale, un campione di tracce orbitali occupa tutto lo spazio disponibile; non si può far variare sistematicamente il parametro di controllo a senza con questo evocare un caos di tipo indesiderato.

L'utente di CHAOS2 specifica quindi un'orbita iniziale battendo sulla tastiera le coordinate di un punto su di essa, poi, comodamente seduto, può guardare affascinato il tracciamento dell'orbita stessa. Potrebbe venir fuori una curva (tracciata non in modo continuo ma intermittente) oppure qualcosa di un po' strano. Per esempio, l'illustrazione in basso della pagina 94 mostra una successione di 38 orbite in una correlazione di Hénon con un valore di 1,111 per a. Dal centro verso l'esterno le orbite formano un insieme annidato di curve chiuse, fino all'improvvisa comparsa di piccole «isole»: orbite individuali incuneate tra quelle più grandi annidate. Più in là, le orbite annidate continuano fino all'avvento del caos. Nelle zone ancora più esterne del diagramma di fase compare un numero maggiore di isole, insieme a una spruzzata casuale di punti che denota l'inizio del caos. Nell'illustrazione una delle aree caotiche, delimitata da un rettangolo, è rappresentata ingrandita. Un avvertimento per chi volesse ingrandire i diagrammi di Hénon: usate l'aritmetica a maggior precisione disponibile sulle vostre macchine.

Come ho ricordato, le correlazioni di Hénon rappresentano una grande varietà di sistemi conservativi, come gli asteroidi orbitanti intorno al Sole. Sfortunatamente, le orbite dei diagrammi non sono le orbite degli asteroidi ma diagrammi di fase di quelle orbite. Nel diagramma appena descritto, l'asse orizzontale può rappresentare la posizione di un asteroide in termini di distanza dal Sole. L'asse verticale può rappresentare la velocità radiale, ossia la velocità di cambiamento di questa distanza. Ciascun punto sull'orbita calcolata dalla rappresentazione di Hénon indica la distanza radiale e la velocità di un asteroide in una specifica posizione angolare rispetto al Sole, cioè quando l'asteroide attraversa un piano verticale posto a quell'angolo con il Sole. Punti successivi calcolati dalla rappresentazione indicano ricomparse successive sul piano. Le isole ricordate in precedenza sono bande di risonanza dovute a perturbazioni nell'orbita dell'asteroide provocate da corpi di maggiori dimensioni presenti nel sistema solare, come Giove. Nelle regioni caotiche, la posizione radiale e la velocità radiale di un asteroide variano in modo sostanzialmente casuale ogni volta che un asteroide torna a visitare il piano specificato. Il suo movimento è imprevedibile: praticamente può accadere quasi qualsiasi cosa.

Dal punto di vista estetico vale la pena di osservare qualche altro tracciato generato da rappresentazioni di Hénon; lasciando da parte le interpretazioni fisiche, un'altra bizzarra caratteristica è ben esemplificata dall'illustrazione riportata in questa stessa pagina: le figure sembrano strane creature acquatiche.

I lettori che vogliano ulteriori informazioni a proposito delle rappresentazioni di Hénon dovrebbero procurarsi il fascicolo di dicembre 1986 della rivista «Byte», dove Gordon Hughes, professore di matematica alla California State University, ha descritto in modo avvincente alcune delle implicazioni fisiche e matematiche delle rappresentazioni di Hénon. Vi si possono trovare anche i listati di alcuni programmi pertinenti, scritti in PASCAL.

Un lettore olandese, Peter de Jong di Leida, ha già suggerito altre formule di iterazione che producono forme e immagini un po' stravaganti. Raccomanda in modo particolare le iterazioni a quattro parametri $x \leftarrow \text{sen}(ay) - \cos(bx)$ e $y \leftarrow \text{sen}(cx) - \cos(dy)$, da iniziare con x e y uguali a 0. Poi, per ottenere la figura che de Jong chiama «zampe di gallina», provate con $a = 2{,}01$, $b = -2{,}53$, $c = 1{,}61$ e $d = 0{,}33$. I valori $-2{,}7$, $-0{,}9$, $-0{,}86$ e $-2{,}2$ generano un «lanciatore di punti» e i valori $-2{,}24$, $0{,}43$, $-0{,}65$ e $-2{,}43$ producono un «uovo di Pasqua ad auto-decorazione».

I lettori sono liberi, come de Jong, di inventarsi le proprie formule di iterazione e di sperimentarle. Chiunque trovi un caos attraente (o sconcertante) è invitato a inviarmelo presso «Scientific American». Crutchfield ha gentilmente acconsentito a corrispondere direttamente con quei lettori di cui non mi è possibile soddisfare i dubbi. Gli si può scrivere al Department of Physics, University of California, Berkeley, California 94720.

I lettori ostinati avranno ormai sicuramente risolto il problema del mese scorso relativo all'inversione della posizione di due vagoni attraverso un ponte poco resistente. I vagoni si trovano su un binario circolare e una locomotiva occupa un altro binario collegato con uno scambio a quello circolare. Il ponte è abbastanza resistente per sopportare un vagone ma non abbastanza per reggere la locomotiva. Come può fare quest'ultima a scambiare di posto i due vagoni?

La locomotiva entra nel percorso circolare, va fino al vagone A e lo spinge sul ponte. Poi torna indietro lungo il binario fino al vagone B, lo aggancia, lo spinge fino al bordo del ponte e aggancia B ad A. Sbuffando, la locomotiva con i due vagoni torna allo scambio ed entra nel binario rettilineo, dove A viene sganciato. Poi riporta B al ponte e ve lo lascia dopo averlo sganciato. Infine, fa il giro del percorso, tira B fuori del ponte portandolo alla sua nuova posizione e va a recuperare A.

Nel fascicolo di giugno sulla musica al calcolatore ho lasciato i lettori a meditare sul modo di ottenere lunghe successioni di note senza ripetizioni scegliendo dei numeri modulo m. Il metodo di scelta richiedeva di partire con un numero seme per poi continuare a moltiplicare per un numero a, sommare un altro numero b e ridurre il risultato prendendo il resto della divisione per m. Se i numeri a e m sono primi uno rispetto all'altro (cioè non hanno in comune fattori maggiori di 1), la successione sarà la più lunga possibile e produrrà anche la musica più strana.

Peter de Jong comunica di aver creato musica strana con il caos. I lettori possono creare suoni analoghi trasformando i numeri generati da CHAOS1 in note musicali. All'esterno delle zone caotiche ci saranno frasi musicali semplici e ripetitive, mentre all'interno ci saranno i veri e propri suoni del caos.

I frattali, la vita e il mondo della natura

Il linguaggio dei frattali

Queste strutture inconcepibilmente ricche di particolari sono molto di più che una semplice curiosità matematica: la geometria frattale è in grado di descrivere in modo compendioso oggetti e processi naturali complessi

di Hartmut Jürgens, Heinz-Otto Peitgen e Dietmar Saupe

«La natura ha fatto uno scherzo ai matematici. Può darsi che ai matematici dell'Ottocento difettasse l'immaginazione, ma alla natura no. Le stesse strutture patologiche che i matematici inventarono per affrancarsi dal naturalismo ottocentesco si rivelarono inerenti agli oggetti familiari da cui siamo circondati.»

FREEMAN DYSON
Characterizing Irregularity
in «Science», 12 maggio 1978

Le «strutture patologiche» ideate dai matematici dell'Ottocento hanno assunto negli ultimi anni la forma di frattali, figure matematiche dotate di dimensione frazionaria e non intera come accade per le ordinarie figure geometriche (per esempio le rette, che hanno dimensione uno, o i piani, che hanno dimensione due). Il fascino che esercitano oggi i frattali è dovuto in ampia misura al lavoro di Benoit B. Mandelbrot, del Thomas J. Watson Research Center della IBM a Yorktown Heights, nello Stato di New York. Mandelbrot coniò il termine frattale nel 1975, traendolo dal latino *fractus*, da *frangere*, cioè «rompere». Fu nel 1983 che il concetto di frattale acquisì vastissima notorietà presso i matematici, gli scienziati e il pubblico non specializzato, con la pubblicazione dell'opera pionieristica *The Fractal Geometry of Nature* dello stesso Mandelbrot.

I frattali sono molto di più che una semplice curiosità matematica: infatti essi offrono un metodo assai conciso per descrivere oggetti e formazioni. Molte strutture hanno una regolarità geometrica soggiacente, detta invarianza rispetto al cambiamento di scala o autosomiglianza. Se si esaminano questi oggetti a scale diverse si incontrano sempre gli stessi elementi fondamentali. Questa configurazione ripetitiva definisce la dimensione frazionaria, o frattale, della struttura. La geometria frattale sembra descrivere le forme e le configurazioni naturali in modo più succinto ed esteticamente più valido rispetto alla geometria euclidea tradizionale.

L'invarianza di scala trova un notevole parallelo nella teoria contemporanea del caos, nella quale molti fenomeni, benché seguano rigide regole deterministiche, si rivelano imprevedibili in linea di principio. Gli eventi caotici, come la turbolenza atmosferica o le pulsazioni cardiache, manifestano andamenti simili su scale temporali diverse, più o meno come gli oggetti dotati di autosomiglianza presentano forme strutturali simili su scale spaziali diverse. La corrispondenza tra frattali e caos non è accidentale; è viceversa il segno di una relazione profonda: la geometria frattale è la geometria del caos.

Un'altra analogia tra geometria frattale e teoria del caos consiste nel fatto che in entrambi i campi le scoperte più recenti sono avvenute grazie alla potenza dei calcolatori moderni. Questi progressi mettono in discussione la visione tradizionale della matematica. Molti matematici hanno salutato l'avvento dei calcolatori come un segno di rinnovamento e di liberazione, ma altri vedono in essi un allontanamento dalla matematica pura.

I frattali sono in primo luogo e soprattutto un linguaggio della geometria; tuttavia i loro elementi fondamentali non possono essere osservati direttamente. Sotto questo profilo essi presentano una differenza basilare rispetto ai ben noti elementi della geometria euclidea, come la retta e il cerchio. I frattali non si esprimono mediante forme primarie, bensì mediante algoritmi, vale a dire insiemi di procedure matematiche. Questi algoritmi vengono poi tradotti in forme geometriche con l'ausilio di un calcolatore. Dal momento che la riserva di elementi algoritmici è inesauribile, quando ci si sia impadroniti del linguaggio frattale si può descrivere la forma di una nube con la stessa precisione e semplicità con cui un architetto può descrivere una casa mediante una pianta tracciata nel linguaggio della geometria tradizionale.

Questa metafora è particolarmente appropriata per le idee fondamentali della geometria frattale. Le lingue indoeuropee sono basate su alfabeti finiti (per esempio le 26 lettere con cui si scrivono le parole inglesi o le 21 lettere della lingua italiana). Le lettere non hanno significato fino a quando non sono giustapposte a formare parole. Analogamente, la geometria euclidea è costituita solo da pochi elementi (la retta, il cerchio e così via), con i quali si possono costruire oggetti complessi che, in un certo senso, solo allora hanno un significato geometrico.

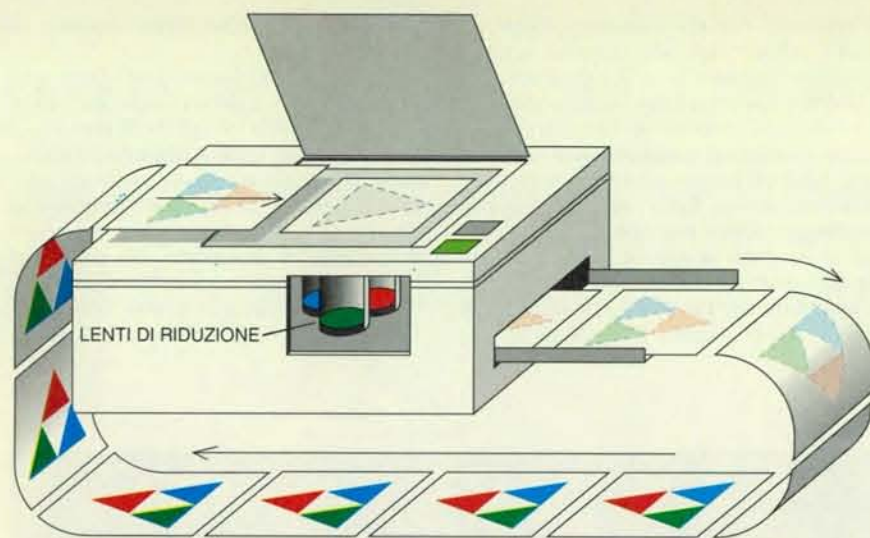
Certe lingue asiatiche come il cinese sono invece costituite da simboli che hanno di per sé un significato. Nel caso di queste lingue il numero di simboli o elementi possibili è arbitrariamente grande e si può considerare infinito. La geometria frattale è costruita più o meno allo stesso modo: è costituita da infiniti elementi, ciascuno dei quali è unico e completo. Gli elementi geometrici sono definiti da algoritmi, che hanno la funzione di unità «semantiche» della lingua frattale.

Le lingue frattali si dividono in due gruppi linguistici principali: quello lineare e quello non lineare. Le lingue di entrambi i gruppi, che «si parlano» usando un numero infinito di algoritmi, contengono un numero infinito di possibili immagini frattali; la lingua dei frattali non lineari, tuttavia, è molto più ricca e varia. In generale i dialetti seguono un insieme deterministico di regole (analoghe alle regole dell'ortografia e della grammatica). Vi è poi una famiglia, quella dei frattali aleatori, che differisce in quanto viene costruita senza regole deterministiche.

Il dialetto fondamentale della lingua frattale è la geometria frattale lineare. Questi frattali sono detti lineari perché i loro algoritmi hanno la stessa forma di quelli che definiscono le rette su un piano (in gergo matematico si dice che contengono solo termini del primo ordine).

Gli algoritmi lineari possono essere studiati con l'ausilio di un immaginario duplicatore di figure: la fotocopiatrice a riduzioni multiple (*si veda l'illustrazione nella pagina seguente in alto*). Si tratta di una metafora del magnifico lavoro di John E. Hutchinson, un matematico della Australian National University a Canberra. Questa macchina funziona più o meno come una normale fotocopiatrice con variatore di riduzione, ma ne differisce per il fatto di avere più lenti di riduzione, ciascuna delle quali può copiare l'originale collocato sulla macchina. Le lenti possono essere predisposte secondo diversi fattori di riduzione e le immagini ridotte possono essere collocate in qualsiasi posizione. La figura può quindi essere spostata, allungata, accorciata, riflessa, ruotata o trasformata in tutti i modi, purché i segmenti di retta dell'originale rimangano segmenti di retta.

Il modo in cui l'immagine viene spostata e ridotta è determinato dall'algoritmo. Mediante un anello di retroazione l'immagine viene elaborata ripetutamente, e tende via via a una forma frattale. Un esempio di frattale ottenuto con un algoritmo ricorsivo (a retroazione) è il triangolo di Sierpinski, dal nome del matematico polacco Waclaw Sierpinski che lo descrisse per primo nel 1916. Questo triangolo è autosimile: ciascuna

Questa immagine al calcolatore è una rappresentazione a tre dimensioni del potenziale elettrostatico che circonda un insieme di Mandelbrot dotato di carica. I valori di potenziale creano un paesaggio fantastico di catene montuose che circondano un vasto lago.

La macchina fotocopiatrice a riduzioni multiple genera una forma frattale tramite un anello a retroazione. Diverse lenti trasformano una immagine iniziale arbitraria (ingresso) in una nuova immagine (uscita), che è un *collage* di copie ridotte dell'immagine iniziale. L'immagine di uscita è poi passata più volte nella macchina dando origine all'immagine finale.

sua parte, per quanto piccola, contiene un'immagine che, ingrandita, riproduce tutto il triangolo.

Il triangolo di Sierpinski può essere costruito con una fotocopiatrice a riduzioni multiple nel modo seguente. Si pone nella macchina un'immagine, la si riduce della metà e la si copia tre volte, una su ciascun vertice di un triangolo equilatero. Il risultato è una configurazione triadica. Poi il procedimento si ripete: questa immagine triadica viene a sua volta ridotta della metà e copiata tre volte, e così via. Dopo che sono state compiute solo sei iterazioni comincia già a delinearsi una forma finale. Questa forma viene chiamata figura limite perché è il risultato al limite di un numero infinito di cicli della fotocopiatrice. La figura limite può essere approssimata in un tempo molto breve, ma non può essere mai ottenuta compiutamente.

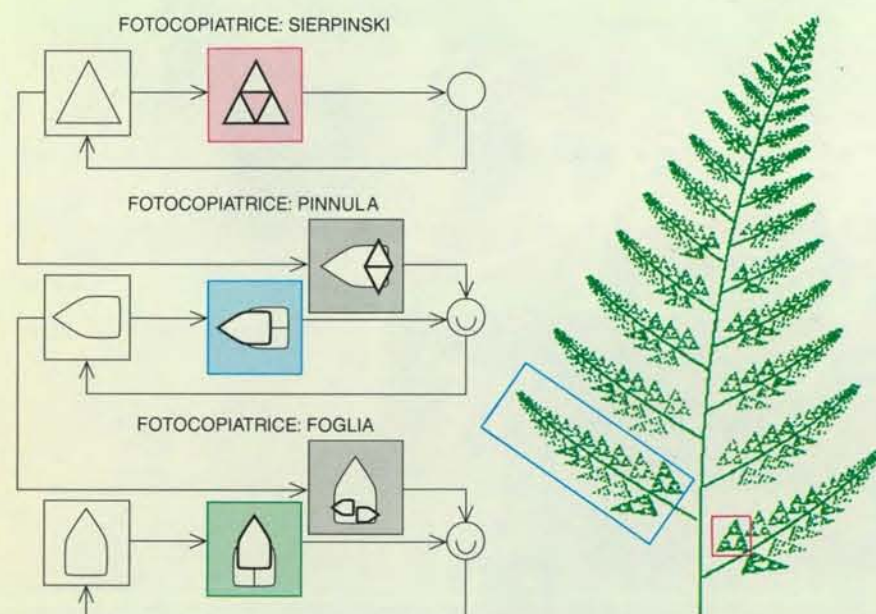
La figura limite non dipende dall'immagine di partenza. Anche se nella fotocopiatrice si colloca un'immagine di partenza molto particolare, come la parola FRACTAL, dopo sei cicli di copiatura la figura di partenza è quasi invisibile ed è la forma del triangolo di Sierpinski a predominare. A ogni copiatura le tracce dell'iniziale parola FRACTAL diventano sempre più indistinte.

Lievi ritocchi alla fotocopiatrice possono dar luogo a figure limite affatto diverse: un albero frattale o un frattale a forma di foglia di felce (*si veda l'illustrazione nella pagina a fronte*). La figura limite dipende solo dalle regole di riduzione e di spostamento (algoritmi) programmate nella macchina.

Queste regole sono casi particolari di quelle che i matematici chiamano trasformazioni lineari affini del piano, vale a dire trasformazioni che conservano la rettilinearità dei segmenti ma ne alterano la scala, la posizione e l'orientazione complessiva. Le regole di un dialetto frattale lineare possono essere descritte completamente da un certo numero (n) di funzioni di trasformazione, indicate con $[f_1, f_2, ..., f_n]$ (*si veda la parte superiore dell'illustrazione nella pagina a fronte*).

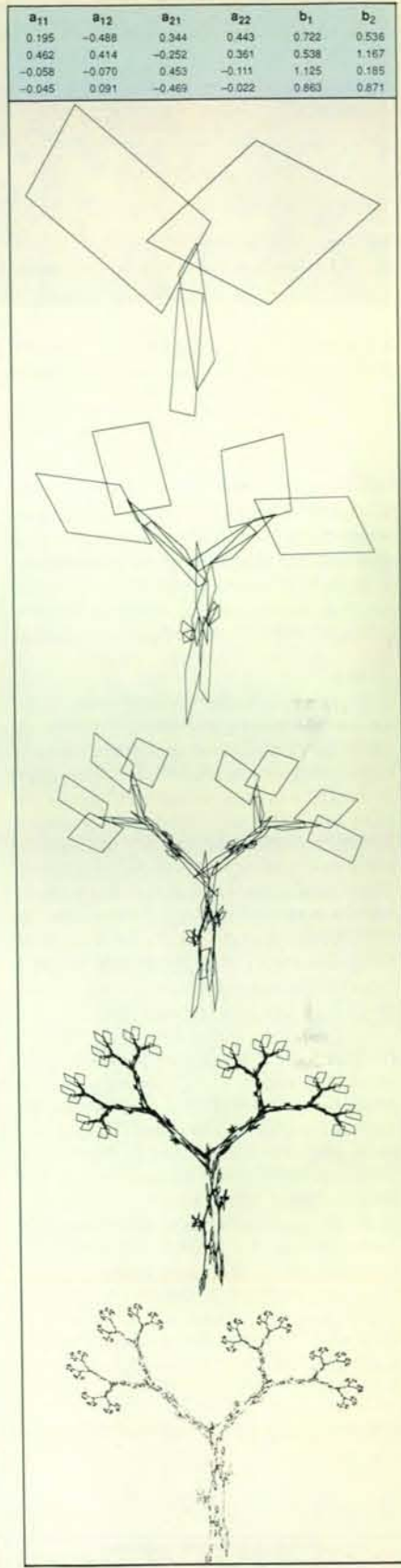
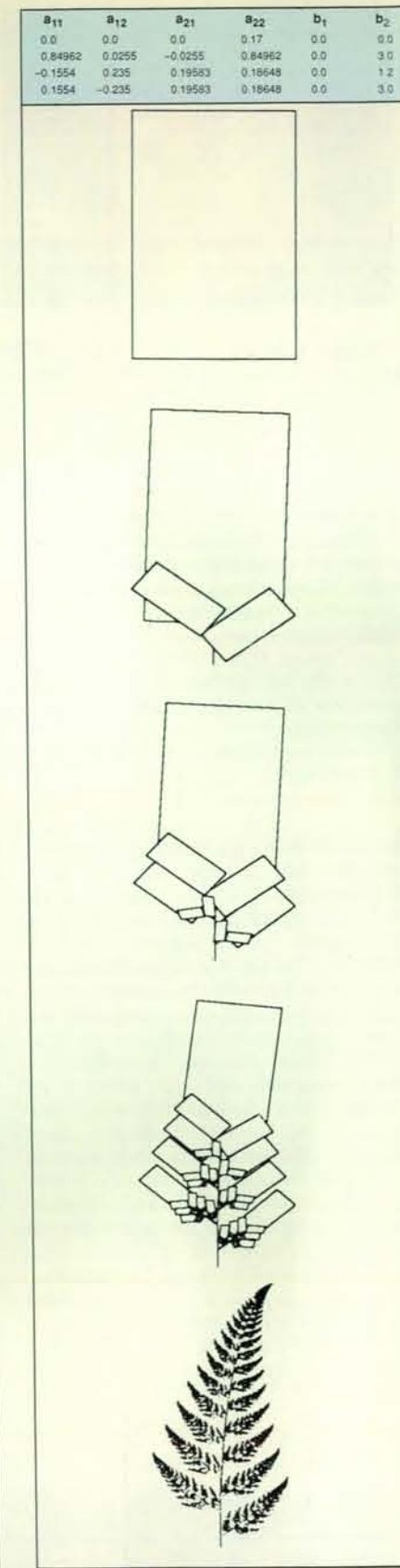
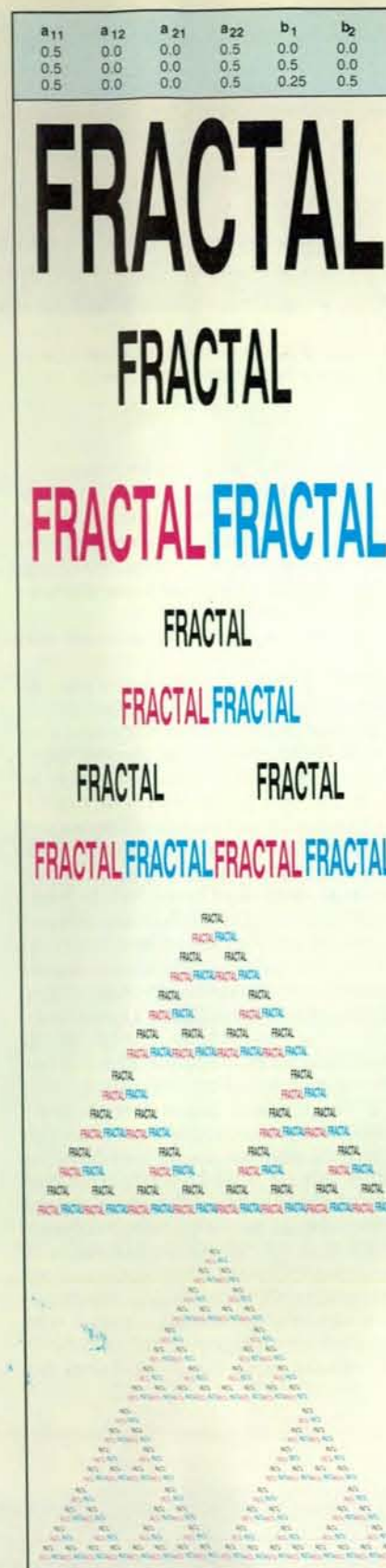
Questo fatto rappresenta una delle grandi potenzialità pratiche della geometria frattale. Descrivendo oggetti opportuni mediante un dialetto frattale lineare si può ridurre notevolmente la quantità di dati necessari per trasmettere o immagazzinare un'immagine. Una dimostrazione convincente di ciò è offerta da una foglia di felce. Una forma così complessa può essere descritta compiutamente da un algoritmo lineare basato su 24 numeri soltanto! Viceversa per rappresentare l'immagine della foglia punto per punto con la qualità di un'immagine televisiva ci vorrebbero parecchie centinaia di migliaia di valori numerici. In linea di principio qualsiasi immagine può essere codificata impiegando l'insieme opportuno di funzioni lineari di trasformazione.

Trasformando le immagini in codici mediante algoritmi frattali si potrebbe ridurre drasticamente il tempo, la complessità e il costo della loro trasmissione via satellite. Questa possibilità solleva un problema cruciale e in gran parte ancora non risolto. Come è possibile ottenere la più piccola famiglia di funzioni di trasformazione $[f_1, ... f_n]$ necessaria a definire un'immagine con una data precisione? Questo problema è attualmente oggetto di molti studi. Tra le applicazioni più generali di questo procedimento si potrebbero menzionare i codici per la costruzione di immagini a mezzatinta o perfino a colori.

La codificazione delle immagini frattali è utile solo se esiste qualche metodo efficiente per estrarre l'immagine imprigionata negli algoritmi frattali. La felce frattale ci fornisce un'opportunità molto interessante di capire come viene prodotta l'immagine. Le regole da fornire

Più macchine fotocopiatrici collegate in una rete possono creare immagini frattali composite, come questa foglia di felce costituita da triangoli di Sierpinski. Si collegano in parallelo più macchine: la prima genera i triangoli di Sierpinski, la seconda li dispone in pinnule e la terza genera la forma totale della foglia (*a sinistra*). Si noti che le pinnule si alternano a destra e a sinistra dello stelo principale; su di esse i triangoli sono contrapposti (*a destra*).

a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
0.5	0.0	0.0	0.5	0.0	0.0
0.5	0.0	0.0	0.5	0.5	0.0
0.5	0.0	0.0	0.5	0.25	0.5

a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
0.0	0.0	0.0	0.17	0.0	0.0
0.84962	0.0255	-0.0255	0.84962	0.0	3.0
-0.1554	0.235	0.19583	0.18648	0.0	1.2
0.1554	-0.235	0.19583	0.18648	0.0	3.0

a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
0.195	-0.488	0.344	0.443	0.722	0.536
0.462	0.414	-0.252	0.361	0.538	1.167
-0.058	-0.070	0.453	-0.111	1.125	0.185
-0.045	0.091	-0.469	-0.022	0.863	0.871

Le immagini frattali generate dall'anello di retroazione della macchina fotocopiatrice dipendono solo dal suo programma di copiatura. La parola FRACTAL viene trasformata da un programma che riduce a metà la grandezza di un'immagine e la copia tre volte, una per ogni vertice di un triangolo equilatero. L'immagine risultante è un triangolo di Sierpinski (*a sinistra*). Trasformazioni analoghe ma un poco più elaborate danno origine a un frattale felciforme (*al centro*) o a un albero frattale (*a destra*). Qualsiasi immagine iniziale introdotta nella fotocopiatrice fornirebbe lo stesso risultato. Sono sufficienti i pochi numeri che definiscono le regole di copiatura (*in alto*) per specificare un'immagine che per essere descritta con metodi tradizionali richiederebbe centinaia di migliaia di numeri.

Gli insiemi di Julia sono frontiere frattali che vengono generate dall'iterazione della trasformazione quadratica $z^2 + c$. La straordinaria varietà delle loro forme dipende solo dal valore del parametro di controllo c. Se il punto c appartiene all'insieme di Mandelbrot, l'insieme di Julia corrispondente è connesso (*nella pagina a fronte*). Gli insiemi di Julia definiti da punti non appartenenti all'insieme di Mandelbrot non sono connessi (*in questa pagina*). L'insieme di Mandelbrot funge da indice degli insiemi di Julia.

alla fotocopiatrice per ottenere questo frattale stabiliscono che ciascuna trasformazione dia luogo a quattro riduzioni e spostamenti dell'immagine precedente. Una delle trasformazioni attua una riduzione molto rapida che schiaccia l'immagine, riducendola a un segmento verticale; e questo segmento costituisce lo «stelo».

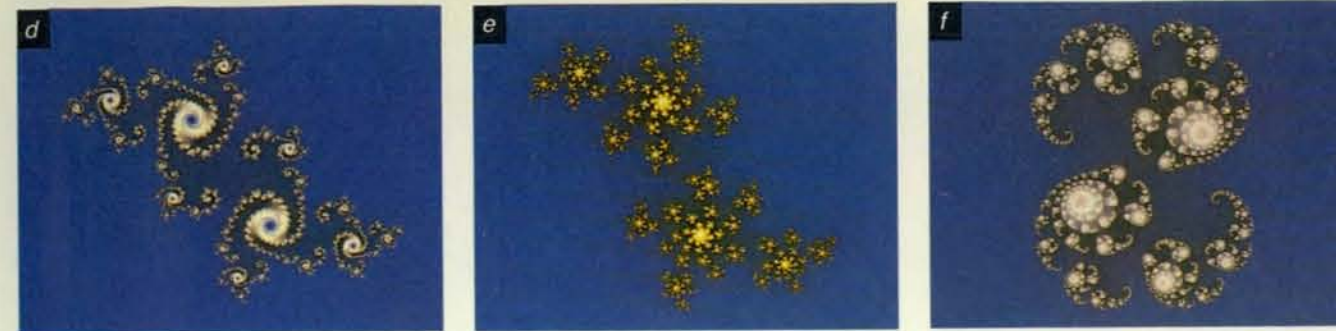
Se si comincia con un solo rettangolo, a ogni copiatura il numero dei rettangoli aumenta di un fattore quattro, e dopo m trasformazioni è pari a 4^m. Dopo quattro iterazioni l'immagine iniziale (in questo caso un rettangolo) può essere ancora facilmente identificata. Un rettangolo abbastanza piccolo da rendere visibile la figura limite (la foglia di felce) si potrebbe ottenere dopo circa 50 iterazioni. Si dovrebbero quindi calcolare e disegnare 4^{50} (più o meno 10^{30}) rettangoli. Nessun calcolatore esistente sarebbe all'altezza di un compito così impegnativo.

Vista questa difficoltà, viene da domandarsi come si possano produrre queste figure limite. L'espediente che consente di ottenere queste immagini è un algoritmo che chiameremo il gioco del caos, proposto da Michael E. Barnsley e Stephen Demko del Georgia Institute of Technology. Il gioco ha inizio con la scelta di un punto arbitrario del piano. Poi viene lanciato un dado a quattro facce, ciascuna delle quali corrisponde a una delle quattro trasformazioni che generano la figura della foglia di felce. Lanciando il dado si individua a caso una delle trasformazioni [f_1, f_2, f_3, f_4], che viene poi applicata al punto scelto e lo trasforma in un nuovo punto del piano. Un altro lancio individua un'altra trasformazione, che viene applicata al punto ottenuto in precedenza e così via. I punti generati dai lanci successivi convergono rapidamente e riempiono densamente la figura limite. Il problema di questa tecnica è che per ottenere la figura limite potrebbe essere necessario un tempo lunghissimo.

Nell'esempio precedente tutte le f_k (dove k rappresenta semplicemente una delle funzioni possibili) hanno la stessa probabilità di essere scelte mediante il lancio del dado. La figura limite può essere ottenuta molto più rapidamente se le probabilità P_k con cui vengono scelte le funzioni f_k nel gioco del caos non sono tutte uguali, cioè se alcune f_k sono più probabili di altre. Il modo più rapido per ottenere la figura limite è quello di assegnare la probabilità più elevata alle funzioni che effettuano la minima riduzione dell'immagine. Con questa modificazione ciascun punto della figura limite è individuato con la stessa frequenza nel gioco del caos, quindi tutte le parti della figura si riempiono con la stessa velocità.

Modificando il gioco del caos si possono descrivere le mezzetinte semplicemente traducendo la frequenza con cui un punto della figura è individuato in un valore della scala dei grigi. Con una scelta opportuna delle P_k, per ogni punto della figura si può ottenere il valore desiderato della scala dei grigi (cioè la frequenza desiderata con cui il punto viene individuato). Applicando questa tecnica ai colori primari additivi (rosso, verde e blu) si possono codificare immagini a colori. In questo modo l'utilità della compressione frattale dei dati viene ulteriormente migliorata.

Per ora non esiste alcun metodo soddisfacente per generare in modo automatico codifiche frattali di una data figura o immagine. Per immagini autosimili come la felce di Barnsley esiste un procedimento semiautomatico che comporta un'interazione tra calcolatore e osservatore. Per prima cosa si scompone l'immagine in parti simili all'immagine complessiva. Nel caso della foglia di felce le due pinnule inferiori sono di forma simile al tutto, così come lo è la parte superiore della foglia che resta quando le pinnule inferiori sono soppresse. Consideriamo una fotocopiatrice a riduzioni multiple contenente trasformazioni che riducano l'immagine intera a queste parti. Ciò si può ottenere facilmente con un procedimento per tentativi ed errori basato su un programma interattivo.

L'idea di base di questo metodo comporta che solo immagini rigorosamente autosimili possano essere codificate in forma frattale. Questa limitazione può essere superata grazie a una promettente estensione del metodo, che è attualmente allo studio. L'idea di fondo è quella di avere molte fotocopiatrici funzionanti simultaneamente in parallelo e organizzate in una rete gerarchica. Questa rete può controllare singoli caratteri autosimili oppure mescolarne più d'uno; ciò consente, per esempio, di creare una foglia felciforme composta di triangoli di Sierpinski (*si veda l'illustrazione a pagina 44 in basso*).

Occupiamoci ora di un altro insieme di dialetti frattali, i dialetti non lineari. Uno di essi, il dialetto quadratico, è stato oggetto di attenzione particolare, poiché produce una grande ricchezza di forme geometriche a partire da un algoritmo piuttosto semplice ed è strettamente collegato all'odierna teoria del caos.

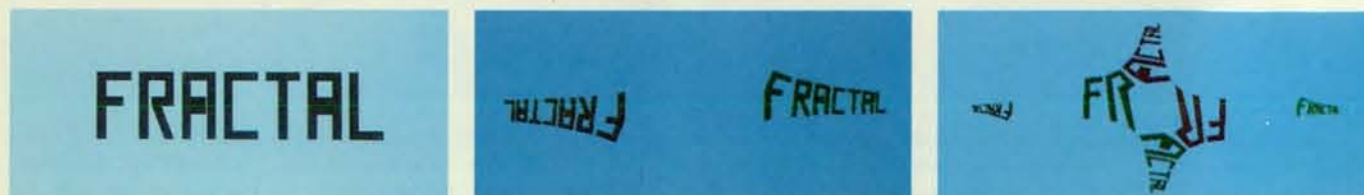
La teoria su cui si basa il dialetto quadratico fu descritta per la prima volta nel 1918 dal matematico francese Gaston Julia, che si trovava allora in un ospedale militare, convalescente delle ferite riportate durante la prima guerra mondiale. Tanto le sue ricerche quanto quelle contemporanee del suo accanito rivale Pierre Fatou furono presto quasi dimenticate, ma di recente il lavoro di Mandelbrot ha riacceso l'interesse per le loro teorie. L'impresa intellettuale di Julia e Fatou è particolarmente notevole perché, non esistendo a quel tempo i calcolatori, essi potevano contare solamente sulle proprie capacità intrinseche di visualizzazione.

Julia e Fatou si occupavano di numeri complessi, ciascuno costituito da un numero reale e da un multiplo di i, l'unità immaginaria definita come la radice quadrata di -1. I numeri complessi vengono di solito rappresentati su un piano con due assi perpendicolari, uno dei quali corrisponde ai numeri reali e l'altro ai numeri immaginari. I due studiosi cercavano di capire che cosa accade a una successione di punti z_k del piano dei numeri complessi generati dalla trasformazione $g(z) = z^2 + c$. Il punto z_{k+1} si ottiene applicando la trasformazione al punto precedente della successione z_k. Il numero complesso c è un parametro di controllo che può essere scelto ad arbitrio. Questo processo iterativo, in apparenza semplice, costituisce la base di una famiglia sbalorditiva di forme.

Quando si applica la trasformazione a un punto iniziale z_0, la successione risultante può comportarsi in due modi diversi: può vagare senza limitazioni, allontanandosi verso l'infinito, oppure restare confinata in una certa regione del piano dei numeri complessi. I punti liberi costituiscono il cosiddetto insieme di fuga; quelli che restano confinati formano il cosiddetto insieme prigioniero. Se il punto di partenza z_0 appartiene all'insieme prigioniero, esso genera una successione che resta in una «prigione» numerica indipendentemente da quante generazioni della successione siano calcolate. La forma della prigione dipende dal valore di c scelto. Per un punto z_0 esterno all'insieme prigioniero, la successione z_k si allontana dal centro del piano e va verso l'infinito. L'insieme prigioniero e l'insieme di fuga sono separati da una frontiera infinitamente stretta, il cosiddetto insieme di Julia (*si veda l'illustrazione in alto in queste due pagine*).

Per quanto appaia sorprendente, l'insieme di Julia si può ottenere anche usando la fotocopiatrice a riduzioni multiple, equipaggiandola con lenti speciali che invertano l'effetto di $g(z)$. L'inversione di $g(z) = z^2 + c$ si compie con due funzioni di trasformazione, $f_1(u) = +(u - c)^{1/2}$ e $f_2(u) = -(u - c)^{1/2}$. (In queste funzioni c è il solito parametro di controllo e u è il valore iniziale scelto.) Queste due funzioni possono essere considerate le «riduzioni» effettuate dalla fotocopiatrice. Facendo funzionare la macchina ripetutamente, punti scelti a caso tendono all'insieme di Julia.

La presenza della radice quadrata nelle equazioni significa che la fotocopiatrice non applica più un fattore di riduzione uniforme. Inoltre, poiché la trasformazione è non lineare, i segmenti di retta vengono trasformati in linee curve. Da un'immagine iniziale emergono due immagini più piccole, poi quattro, poi otto, finché pian piano si forma la figura limite (*si veda l'illustrazione in basso in queste due pagine*). Come nel caso dei frattali lineari, la figura limite non dipende dalla particolare immagine di partenza, ma è completamente determinata da f_1 e f_2, cioè dalla scelta del parametro c.

Si presenta ora uno dei problemi più difficili e affascinanti della geometria frattale. Tornando alla metafora della lingua, il problema può essere tradotto in questa domanda: quali sono le regole grammaticali dei dialetti quadratici? In termini matematici il problema è: l'infinita varietà degli insiemi di Julia ammette un principio ordinatore soggiacente?

Anche i frattali non lineari come gli insiemi di Julia possono essere generati tramite una fotocopiatrice a riduzioni multiple opportunamente programmata. La macchina ora non effettua più semplici operazioni di riduzione, ma curva e distorce l'immagine di partenza oltre a rimpicciolirla. La fotocopiatrice compie per via grafica un'inversione delle equazioni quadratiche che definiscono un insieme di Julia effettuando le trasformazioni $+(z - c)^{1/2}$ e $-(z - c)^{1/2}$. La figura limite generata dalla fotocopiatrice è un insieme di Julia.

Nel cercare la risposta si è giunti a una delle più belle scoperte della matematica sperimentale. La soluzione sta nel fatto, noto a Julia e a Fatou, che per ogni parametro di controllo c l'immagine frattale risultante è di due tipi possibili: l'insieme di Julia può essere un unico insieme connesso, oppure essere costituito da un numero infinito di punti non connessi, come polvere.

Supponiamo di tracciare un punto per ogni parametro di controllo c del piano complesso appartenente a un insieme di Julia connesso e di lasciare uno spazio vuoto per ogni c appartenente a un insieme di Julia non connesso. Il risultato è l'insieme, ormai celebre, di Mandelbrot, un frattale dalla ricchezza davvero straordinaria.

Naturalmente per decidere se un punto c appartiene all'insieme di Mandelbrot, è necessario sapere se un dato insieme di Julia è connesso. Uno dei grandi successi di Julia e Fatou fu la scoperta che questo difficile problema può essere risolto con un semplice calcolo. Si consideri la successione di valori z_k generata dalla funzione $g(z) = z^2 + c$ quando il punto iniziale z_0 coincide con lo zero. In tal modo l'attenzione si concentra sul fattore cruciale, il parametro di controllo c. La successione risultante è $0, c, c^2 + c, (c^2 + c)^2 + c, \ldots$. Se questa successione non diverge verso l'infinito allora il corrispondente insieme di Julia è connesso e il punto c appartiene all'insieme di Mandelbrot.

Ogni porzione dell'insieme di Mandelbrot caratterizza una famiglia di insiemi di Julia correlati. Per esempio il corpo principale, a cardioide, dell'insieme di Mandelbrot caratterizza gli insiemi di Julia che somigliano a cerchi raggrinziti. Benché l'insieme di Mandelbrot non sia esattamente autosimile come il triangolo di Sierpinski e la foglia di felce frattale, esso possiede una proprietà analoga: se si ingrandisce la sua frontiera, l'insieme di Mandelbrot rivela un numero infinito di minuscole copie di se stesso. La ricchezza di forme e di strutture contenute nell'insieme di Mandelbrot può essere apprezzata solo quando lo si esamini con grande minuzia.

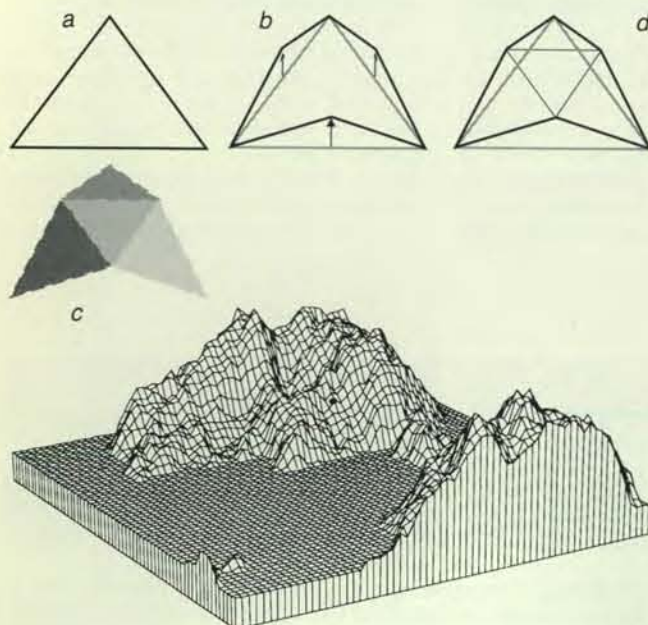
La proprietà forse più affascinante dell'insieme di Mandelbrot è che esso può essere considerato un «deposito» di immagini di efficienza infinita: oltre a suddividere gli insiemi di Julia in connessi e non connessi, l'insieme di Mandelbrot funge anche da indice diretto e grafico di un numero infinito di insiemi di Julia. Ingrandendo l'insieme di Mandelbrot intorno a un punto c situato sulla sua frontiera, appaiono forme che sono anche gli elementi costitutivi dell'insieme di Julia corrispondente al punto c. Questa scoperta, tuttavia, non è stata ancora rivestita di tutto il necessario rigore matematico. Tan Lei, un giovane ricercatore di talento che lavora all'Università di Lione, ha dimostrato che l'insieme di Mandelbrot si comporta in questo modo per la maggior parte dei valori del parametro c situati esattamente sulla frontiera dell'insieme.

Le proprietà dell'insieme di Mandelbrot sono state e sono tuttora un gran de cimento per la ricerca matematica. Enormi progressi sono stati compiuti grazie alla fusione di teoria matematica ed esperimenti di eidomatica (ossia di grafica al calcolatore), in particolare nei fondamentali lavori di Adrien Douady dell'École Normale Supérieure di Parigi e di John H. Hubbard della Cornell University.

Il lavoro di gran lunga più riuscito in questo campo è quello sul cosiddetto potenziale elettrostatico dell'insieme di Mandelbrot. Si immagini che l'insieme sia dotato di carica elettrica. Si potrebbe misurare il potenziale collocando una carica puntiforme all'esterno dell'insieme e misurando la forza elettrostatica agente su quel punto. Risulta che il calcolo del potenziale è strettamente legato alla serie $0, c, c^2 + c, (c^2 + c)^2 + c \ldots$, usata per stabilire se un punto c appartiene o no all'insieme di Mandelbrot.

Si è visto che generare una rappresentazione tridimensionale del potenziale è laborioso, specie nelle animazioni usate per studiare l'insieme di Mandelbrot. Esaminando meglio le proprietà eidomatiche del potenziale, di recente si è potuto ridurre di un ordine di grandezza il costo in tempo macchina. Di conseguenza molti ricercatori, noi compresi, studiano sempre più l'insieme di Mandelbrot mediante video animati dal calcolatore. Ricerche analoghe sono in corso anche su rappresentazioni tridimensionali del potenziale di altri frattali.

Tutti i frattali finora esaminati possono essere considerati deterministici. Benché i processi aleatori (per esempio il lancio di un dado) possano aiutarci a produrre immagini frattali, essi non hanno alcun effetto sulla forma frattale finale. La situazione è affatto diversa per un'altra classe di frattali, i cosiddetti frattali aleatori.

Per generare un frattale di questo tipo si può cominciare con un triangolo giacente su un piano arbitrario. I punti medi di ciascun lato del triangolo vengono collegati tra loro e il triangolo viene così diviso in quattro triangoli più piccoli. Ciascun punto medio viene poi alzato o abbassato di una quantità scelta a caso. Lo stesso procedimento viene applicato a ciascuno dei triangoli più piccoli e il processo viene ripetuto all'infinito. All'aumentare del numero delle iterazioni, comincia a formarsi una superficie sempre più ricca di particolari.

Si possono creare paesaggi frattali con il metodo dello spostamento dei punti medi. I punti medi dei lati di un triangolo (*a*) vengono uniti da segmenti e spostati in su o in giù, fuori dal piano dell'immagine (*b*). Si ottengono così quattro piccoli triangoli su cui si ripete il procedimento. Una legge di distribuzione stabilisce l'entità dello spostamento e quindi determina la scabrosità del terreno frattale. Un programma eidomatico genera poi ombreggiature appropriate (*c*), dando vita a risultati straordinariamente realistici (*d*).

L'insieme di Mandelbrot rispecchia l'ordine soggiacente alla infinita varietà degli insiemi di Julia. Tutti i suoi punti rappresentano valori del parametro c corrispondenti a insiemi di Julia connessi. Se il punto c non appartiene all'insieme di Mandelbrot l'insieme di Julia a esso associato non è connesso. L'insieme di Mandelbrot contiene una ricchezza di dettagli inimmaginabile. Tre ingrandimenti successivi dell'insieme rivelano strutture simili che si ripetono, fra cui anche copie in miniatura dell'insieme stesso, oltre a forme nuove e differenti. Se l'intero insieme fosse raffigurato alla scala dell'immagine di destra, sarebbe grande come 100 campi da calcio.

In questo «metodo dello spostamento dei punti medi», l'entità aleatoria dello spostamento dei punti medi è retta da una legge di distribuzione che può essere modificata fino a ottenere una buona approssimazione della superficie di cui si vuol costruire il modello. Per un modello di una superficie relativamente liscia, le trasformazioni usate dovrebbero prevedere una regola per cui gli spostamenti dei punti medi diventino piccolissimi già dopo poche iterazioni. Una regola del genere aggiunge solo piccole prominenze sullo sviluppo complessivo. Per rappresentare invece una superficie accidentata, per esempio la topografia di una catena montuosa, è meglio far diminuire di poco l'entità degli spostamenti a ogni iterazione.

Questo metodo per costruire superfici ha molte applicazioni. È stato impiegato per ottenere modelli dell'erosione del suolo e per analizzare le registrazioni sismiche al fine di capire i cambiamenti nelle zone di faglia. Questo concetto è stato usato da Richard F. Voss, collega di Mandelbrot al Thomas J. Watson Research Center, per generare immagini molto realistiche di pianeti, satelliti, nubi e montagne (*si veda l'illustrazione nella pagina a fronte*).

A prescindere dalla loro origine e dal metodo di costruzione, tutti i frattali presentano una caratteristica importante: se ne può misurare la scabrosità, la complessità o l'accartocciamento mediante un numero caratteristico, la dimensione frattale. Le varie definizioni concettuali della dimensione frattale risalgono più o meno a un lavoro del 1919 del matematico Felix Hausdorff dell'Università di Bonn.

La dimensione frattale può essere determinata mediante un procedimento di conteggio proposto da Mandelbrot. Si consideri una forma complessa a cui è sovrapposto un reticolo di quadrati tracciati su carta millimetrata. Alcuni quadrati conterranno parte della forma, altri saranno vuoti. Il numero N di quadrati non vuoti dipende dalla forma data e dal lato E dei quadrati del reticolo. Si postula che N sia proporzionale a $1/E^D$ (più è fitto il reticolo più sono i quadrati non vuoti). L'esponente D è la dimensione. Per una figura piana, per esempio un cerchio, quando il lato dei quadrati del reticolo viene ridotto a metà il numero dei quadrati non vuoti dovrebbe essere moltiplicato per quattro (due al quadrato), perché la figura ha dimensione due. Per un frattale il numero dei quadrati non vuoti verrebbe moltiplicato per un valore frazionario un po' più grande o un po' più piccolo.

Questo procedimento non è limitato alle forme o agli oggetti matematici contenuti in un piano: si può anche calcolare la dimensione frattale di cose reali come fiumi, nubi, litorali, alberi, arterie o villi intestinali. Le arterie umane, per esempio, hanno una dimensione frattale di circa 2,7.

Oltre a essere utile per descrivere la complessità degli oggetti naturali, la geometria frattale offre un'interessante possibilità per rinnovare l'insegnamento della matematica. I concetti della geometria frattale sono evidenti e intuitivi e le forme che s'incontrano possiedono una grande attrattiva estetica e un'ampia gamma di applicazioni. La geometria frattale può quindi contribuire a sfatare l'idea che la matematica sia arida e inaccessibile e può motivare gli studenti ad apprendere questa sconcertante ed eccitante disciplina.

Di fronte al linguaggio dei frattali, così nuovo e in evoluzione tanto rapida, gli stessi scienziati e matematici provano una sorta di stupore infantile. Scrive Mandelbrot:

«Gli scienziati... saranno sorpresi e compiaciuti nello scoprire che non poche forme che prima dovevano chiamare *granulose, tentacolari, intermedie, foruncolose, butterate, ramificate, algali, strane, intricate, tortuose, serpeggianti, esili, grinzose* e così via potranno d'ora in poi essere studiate in modo rigorosamente e vigorosamente quantitativo.

«I matematici... saranno sorpresi e compiaciuti nello scoprire che gli insiemi [frattali] considerati finora eccezionali... costituiscono in un certo senso la regola, che certe costruzioni ritenute patologiche discendono in modo naturale da problemi molto concreti e che lo studio della natura può contribuire a risolvere problemi antichi e a trovarne tanti di nuovi.»

BIBLIOGRAFIA

MANDELBROT B. B., *The Fractal Geometry of Nature*, W. H. Freeman & Co., 1983.

PEITGEN H. O. e RICHTER P., *La bellezza dei frattali*, Bollati Boringhieri, Torino, 1987.

BARNSLEY M., *Fractals Everywhere*, Academic Press, Inc., 1988.

PEITGEN H. O. e SAUPE D., *The Science of Fractal Images*, Springer-Verlag, 1988.

PEITGEN HEINZ-OTTO, JÜRGENS HARTMUT e SAUPE DIETMAR, *Fractals for the Classroom*, Springer-Verlag, 1989.

PEITGEN H. O., JÜRGENS H., SAUPE D. e ZAHLTEN C., *Fractals: An Animated Discussion*, con Edward Lorenz e Benoit B. Mandelbrot (video), W. H. Freeman & Co., 1990.

L'accrescimento dei frattali

In natura esistono strutture con particolari schemi di ramificazione, dette frattali, il cui meccanismo di crescita può illustrare fenomeni come la cristallizzazione di certi solidi o il movimento delle bolle d'aria nei fluidi

di Leonard M. Sander

Lo studio macroscopico della struttura della materia pone problemi di grande difficoltà. Ogni pezzetto di materia contiene un numero enorme di atomi e molecole, spesso disposti in maniera complicata e disordinata. Nel caso di un cristallo perfetto o di un liquido in movimento laminare, la struttura è uniforme su grande scala. Ma la maggior parte dei fenomeni naturali complessi, come il moto turbolento di un fluido, l'accrescimento delle particelle metalliche in un bagno elettrolitico o la formazione delle catene montuose, sfida la nostra ragione.

Negli ultimi dieci anni sono comunque stati fatti passi importanti verso la comprensione di tali fenomeni, utilizzando spesso il concetto rivoluzionario di frattale, un termine proposto da Benoit B. Mandelbrot del Thomas J. Watson Research Center della IBM a Yorktown Heights, nello Stato di New York. Un frattale è un oggetto con una complessa struttura sottilmente ramificata. Ingrandendo gradualmente una parte della struttura, vengono alla luce dettagli che si ripetono identici a tutte le scale d'ingrandimento. Un frattale appare quindi identico se lo si osserva a grande, piccola o piccolissima scala, per esempio con lato di un metro, un millimetro o un micrometro (un milionesimo di metro). Mandelbrot ha osservato che molti oggetti naturali, apparentemente disordinati, godono di questa proprietà.

Un crescente numero di osservazioni dimostra che le forme frattali sono diffusissime in natura. Frattali conosciuti come aggregati di percolazione corrispondono molto bene all'andamento di un fluido che filtra attraverso una matrice solida, come l'acqua nel terreno o il caffè attraverso i grani macinati. Il nerofumo, i colloidi e alcuni polimeri paiono avere struttura frattale. Frattali si incontrano anche nel movimento di bolle d'aria nell'olio, nella crescita di alcuni cristalli e nella struttura di scariche elettriche del tipo dei fulmini. Anche le forme casuali delle nuvole e delle linee costiere sono quasi certamente dei frattali.

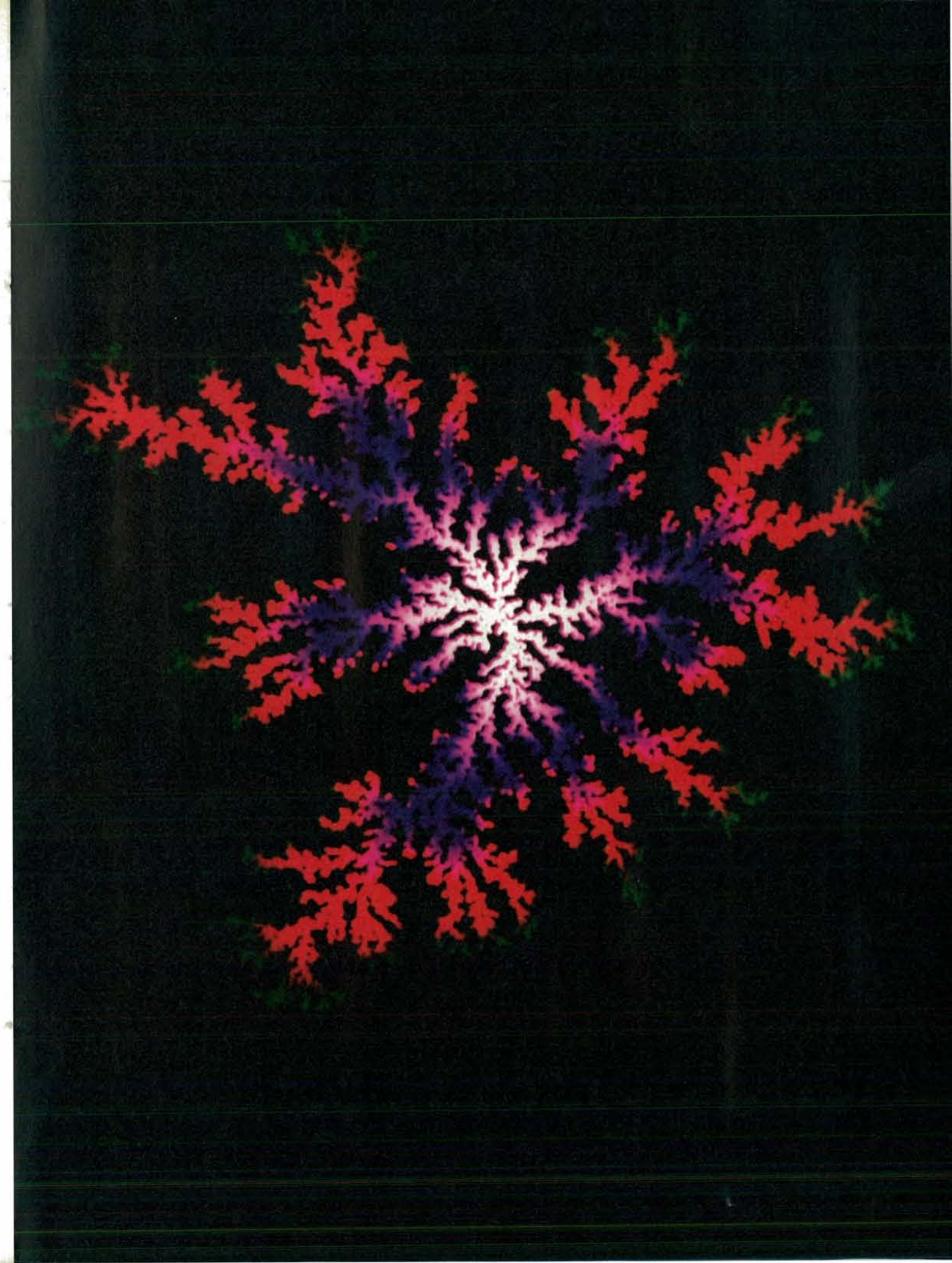
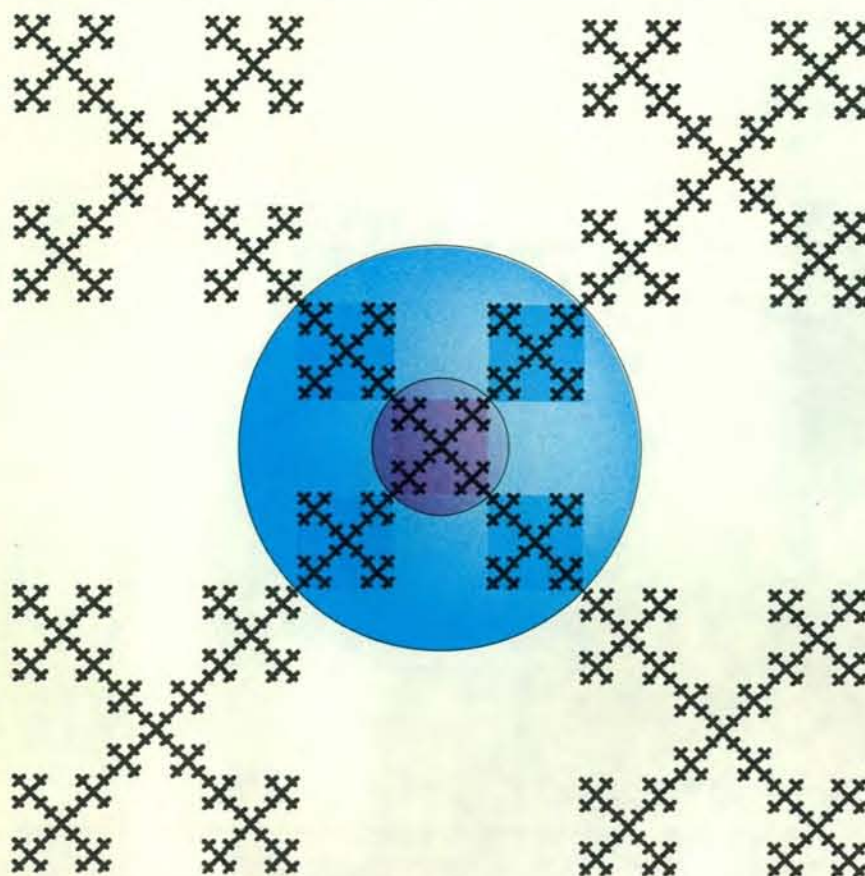
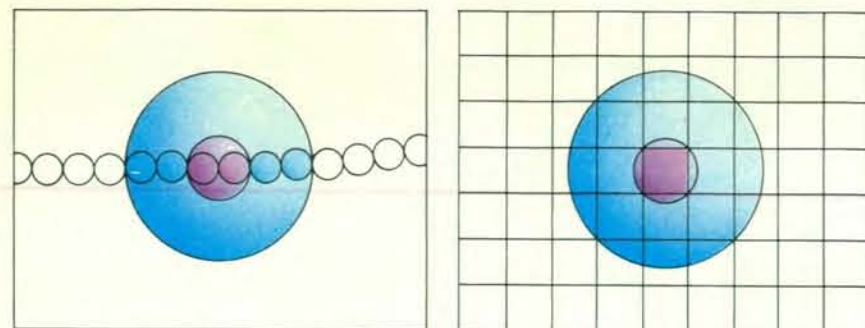
Con l'aumentare delle osservazioni riguardanti l'esistenza dei frattali in natura, si è iniziato a studiare come possono formarsi simili strutture. Thomas Witten, Jr., della Exxon Research and Engineering Company e io abbiamo proposto nel 1981 un meccanismo di accrescimento dei frattali che abbiamo chiamato aggregazione per diffusione. Secondo il nostro modello, da un processo di accrescimento disordinato e irreversibile può avere origine una struttura frattale. La teoria è interessante per due motivi. Innanzitutto è concettualmente semplice e consente di realizzare facilmente dei modelli al calcolatore. E poi, cosa più importante, pare spiegare come si formano in natura diversi tipi di frattali.

Quali sono le proprietà dei frattali? Studi astratti su oggetti che ora sappiamo essere dei frattali sono stati condotti, molto prima di Mandelbrot, da altri matematici che li consideravano delle specie di mostri di interesse puramente accademico. Eppure un tipico frattale assomiglia più a un fiocco di neve che a un mostro. La ragione di ciò sta nella ripetizione continua di una struttura elementare. Per esempio, ogni unità del frattale mostrato nella parte in basso dell'illustrazione a pagina 70 è costituita da cinque subunità identiche. Cinque unità più grandi possono poi aggregarsi per farne una più grande ancora e così via. Ogni generazione contiene dei vuoti di dimensioni legate alla scala di quella particolare generazione. La struttura è invariante anche rispetto alla scala: a ciascuno stadio di crescita, qualsiasi parte con diametro pari a un terzo di quello dell'intera figura è del tutto identica a questa. L'invarianza rispetto alla scala è una «simmetria» dei frattali. Come gli oggetti con sezione circolare sono invarianti rispetto alla rotazione, i frattali sono invarianti rispetto alla dilatazione, ovvero al cambiamento di scala.

È utile avere una misura del modo in cui un frattale è invariante rispetto ai cambiamenti di scala. Questa è data da un numero denominato dimensione frattale. Le dimensioni frattali, al contrario delle dimensioni ordinarie, non sono espresse con un intero bensì con una frazione. Il frattale di cui stiamo parlando ha, per esempio, dimensione 1,46. Esso sta a metà fra una linea retta monodimensionale e un piano bidimensionale. Quanto più un frattale riempie un piano, tanto più la sua dimensione si avvicina a 2. La struttura dell'illustrazione nella pagina a fronte, ottenuta al calcolatore simulando il processo di aggregazione per diffusione, ha dimensione frattale 1,71. Una proprietà di questo frattale - ma, in effetti, di tutti i frattali - è che la sua densità diminuisce con l'accrescimento.

La dimensione frattale di un oggetto fisico è una sua proprietà «universale», cioè indipendente da molti dettagli riguardanti il modo nel quale si è formato. La dimensione frattale, come altre proprietà universali, è correlata con il comportamento a grande scala, dove i particolari della struttura si compensano ampiamente nella media. Ne segue che un modello semplificato, che trascuri gran parte della complessità di un sistema reale, può comunque descriverne correttamente le proprietà di scala.

L'importanza del modello di aggrega-

La struttura frattale è stata prodotta da Paul Meakin della E. I. du Pont de Nemours & Company, Inc., simulando al calcolatore un processo di aggregazione per diffusione. 50 000 «particelle» sono state liberate, una alla volta, da una zona fuori della figura, lasciandole diffondere verso il centro fino a venire in contatto l'una con l'altra e formare un aggregato che si accresce. I colori indicano l'ordine di arrivo delle particelle: quelle bianche sono le prime e quelle verdi le ultime.

zione per diffusione sta nel fatto che mostra una correlazione tra i frattali e i processi di accrescimento. Gli oggetti naturali possono crescere in molti modi diversi. Un cristallo perfetto, per esempio, cresce in condizioni prossime all'equilibrio: dopo aver «tentato» diverse configurazioni perviene a quella più stabile. Una molecola che va ad aggiungersi al cristallo deve, in generale, cercare una posizione adatta fra le molte possibili. Il cristallo si forma quindi lentamente e le sue superfici sono soggette a un continuo riarrangiamento. Tuttavia, la maggior parte dei processi che avvengono in natura non hanno la possibilità di raggiungere l'equilibrio. I fenomeni biologici, per esempio, sono lontani dall'equilibrio e ciò vale anche per i frattali che intendiamo descrivere (alcuni frattali crescono in prossimità dell'equilibrio, ma non ne parleremo in questo contesto).

La dimensione frattale differisce dalle dimensioni ordinarie poiché non viene espressa con un numero intero. Per determinare la dimensione frattale di una struttura, si conta il numero medio N di unità ripetitive fondamentali presenti entro una sfera di raggio r centrata in un punto qualsiasi. Secondo la geometria euclidea, il numero di unità fondamentali è pari a una costante C moltiplicata per il raggio elevato al valore della dimensione D ($N = C \times r^D$). Nel caso di una linea la dimensione è, ovviamente, 1. Triplicando il raggio della sfera si triplica il numero delle unità in essa contenute (*in alto a sinistra*). Per la materia ordinaria triplicando il raggio della sfera si moltiplica per 9 il numero delle unità (*in alto a destra*). In un frattale (*in basso*) di dimensioni 1,46, invece, triplicando il raggio, il numero delle unità cresce di 5 volte. Cresce quindi più velocemente che in una linea, ma non come su un piano di materia ordinaria; quindi, si tratta di un oggetto intermedio fra una linea e un piano. Il frattale dell'illustrazione è stato ideato da Tamás Vicsek dell'Istituto di fisica tecnica dell'Accademia delle scienze di Budapest.

Immaginiamo di far crescere un aggregato di particelle aggiungendone una alla volta, in modo che appena viene a contatto con l'oggetto in accrescimento vi aderisca, senza cercare altri siti di legame. Questo processo viene detto di aggregazione. È un esempio estremo di processo lontano dall'equilibrio perché non vi avviene alcun riarrangiamento. Ora, supponiamo che le particelle diffondano verso l'aggregato seguendo un cammino casuale, costituito da una successione di passi di lunghezza e direzione casuali. Un cammino monodimensionale casuale potrebbe essere realizzato da una persona che lanci in aria una moneta e faccia un passo in avanti o all'indietro a seconda che sia uscita testa o croce. Il processo di aggregazione di particelle che si muovono secondo un cammino casuale è appunto quello che Witten e io chiamiamo aggregazione per diffusione.

È possibile far crescere piccoli aggregati anche su un calcolatore personale. Si comincia con il porre una particella in corrispondenza dell'origine dell'aggregato. Poi si libera un'altra particella a una qualche distanza dalla prima e la si lascia migrare seguendo un cammino casuale fino a raggiungere e toccare la prima. A questo punto si libera una terza particella lontano dall'aggregato e così via. La simulazione mostra che gli aggregati prodotti con questo procedimento sono dei frattali.

Malgrado l'aggregazione per diffusione sia semplice da descrivere e da simulare, il processo non è ancora ben chiaro. Perché, per esempio, vengono prodotti dei frattali e non dei complessi amorfi e privi di simmetria? Perché si formano raramente degli anelli? Quali sono le relazioni di dipendenza della dimensione frattale dalla dimensione dello spazio nel quale avviene l'aggregazione? Le risposte a simili quesiti sono ancora lontane e pongono un problema notevole ai fisici teorici, perché sembra che in questo caso i consueti strumenti matematici non funzionino più.

Tuttavia, alcune caratteristiche del processo possono essere comprese in maniera qualitativa. Immaginiamo di iniziare facendo depositare particelle su un aggregato con la superficie liscia. Quando l'aggregato è piccolo più particelle aderiscono per puro caso in una certa regione. Si formano in tal modo sulla sua superficie piccole protuberanze e buche, in conseguenza del «rumore» imputabile al comportamento casuale delle particelle.

Una volta formatesi, le protuberanze cresceranno più rapidamente delle buche. Questo perché una particella che si muove casualmente verso l'aggregato ha una probabilità elevata di aderire nei pressi della sommità di una protuberanza o sui suoi fianchi prima di cadere entro una buca. Dato che la crescita avviene prevalentemente presso la sommità delle protuberanze, queste diventano sempre più alte e il riempimento delle

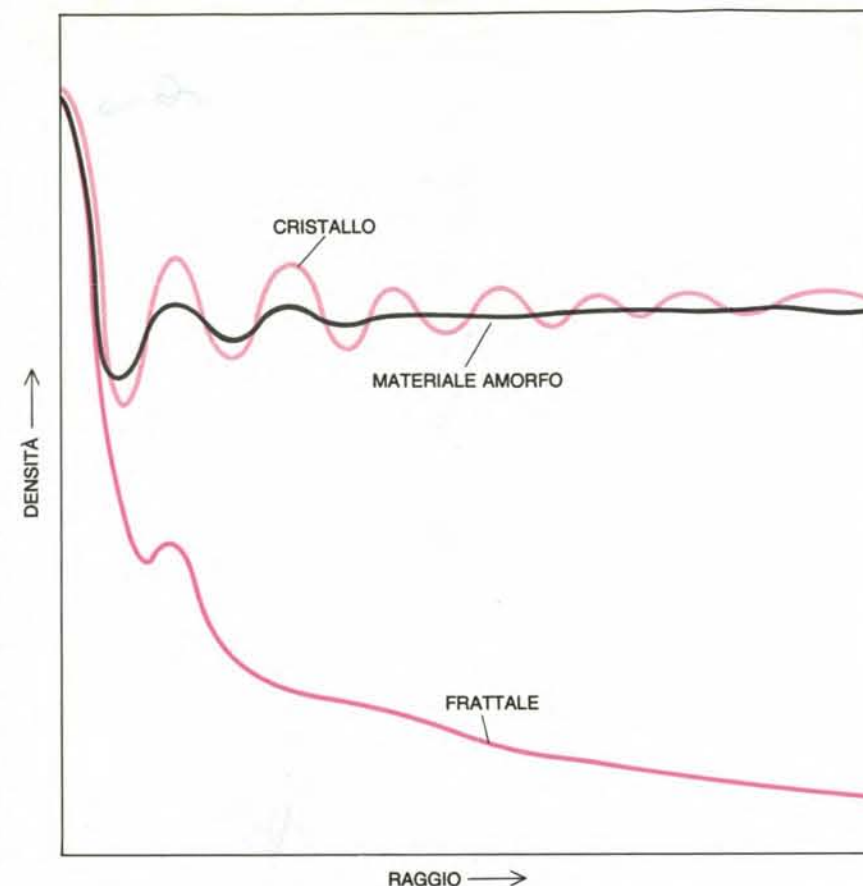
buche diviene sempre meno probabile. Ne segue che l'aggregato, inizialmente appena distorto, cresce sempre più distorto, effetto che viene detto instabilità di accrescimento. È presumibile che alla fine la crescita e la suddivisione delle sommità delle protuberanze diano origine a un frattale. Malgrado i particolari rimangano ignoti, è certo che l'origine della complessità degli aggregati cresciuti per diffusione va cercata nell'interazione fra rumore e crescita.

La densità di un frattale diminuisce con l'aumentare delle sue dimensioni, mentre la densità di un solido cristallino o di un qualunque materiale amorfo tende a un valore costante.

Negli ultimi cinque anni l'aggregazione per diffusione è stata oggetto di ampi studi. Molto entusiasmo deriva dal fatto che il modello pare in grado di descrivere eventi reali. Accade infatti spesso che particelle migrino fino a fissarsi in un sito. Robert M. Brady e Robin C. Ball dell'Università di Cambridge hanno segnalato nel 1984 che l'aggregazione per diffusione è, per esempio, una buona idealizzazione di ciò che accade nella deposizione di un metallo da una soluzione elettrolitica di ioni che diffondono. Malgrado il meccanismo dell'adesione degli ioni al deposito sia diverso, nei particolari, da ciò che avviene nella simulazione al calcolatore, la circostanza non pare avere effetto sulla struttura globale che ne risulta. E tali particolari non hanno importanza neppure per quanto riguarda la dimensione frattale.

Per esempio, il deposito di zinco metallico in una cella elettrolitica (*si veda l'illustrazione in alto a sinistra a pagina 76*) somiglia moltissimo al frattale generato al calcolatore nell'illustrazione a pagina 69. Il deposito di zinco ha una dimensione frattale pari a 1,7 che, tenendo conto dell'errore sperimentale, concorda con la dimensione del frattale generato al calcolatore (1,71). Questa concordanza è un esempio notevole di universalità e di invarianza di scala: nella simulazione al calcolatore sono stati impiegati 50 000 punti, mentre il numero di atomi di zinco nel deposito è enorme, almeno un miliardo di miliardi.

In effetti, alterando le regole nella simulazione al calcolatore, si mettono in luce diversi tipi di universalità. Poniamo, per esempio, che a volte la particella rimbalzi via dall'aggregato, anziché aderirvi. Questa regola, che è una semplice rappresentazione di una delle complicazioni che possono verificarsi nella realtà, porta a un ispessimento dei rami dell'aggregato, ma non ne cambia la dimensione frattale.

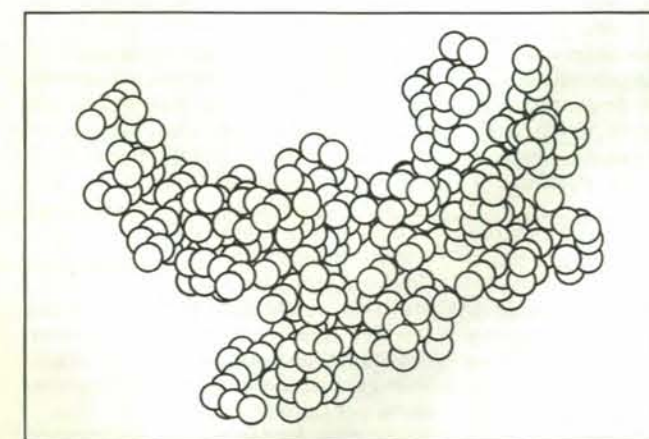
Certo è facile credere che la deposizione di ioni metallici su un elettrodo possa venire descritta come un'aggregazione per diffusione. È però rilevante il fatto che il modello riesca a descrivere molti altri fenomeni. Uno di questi implica un apparecchio chiamato cella di Hele-Shaw, derivato dagli studi di un ingegnere navale inglese del secolo scorso, Henry S. Hele-Shaw. La cella è costituita da due lastre di vetro parallele, fra le quali è interposto un liquido viscoso, come la glicerina. Quando si inietta nel centro della cella un fluido meno viscoso, per esempio aria, la glicerina viene spostata e si forma una bolla d'aria, dalla quale sporgono svariate protuberanze, o «dita» (*si veda l'illustrazione in alto a destra a pagina 76*). Il fenomeno viene giustamente chiamato ramificazione digitiforme viscosa o, più brevemente, ramificazione viscosa (*viscous fingering*); esso ha grande importanza pratica perché si verifica quando si inietta acqua nel centro di un campo petrolifero per migliorare il recupero spingendo il petrolio verso i pozzi periferici. In conseguenza della ramificazione viscosa, molto petrolio rimane intrappolato a meno che vengano impiegate tecniche particolari.

La struttura delle ramificazioni viscose mostra una forte somiglianza con le immagini di aggregati per diffusione ottenute al calcolatore. La spiegazione di tale fatto è stata data recentemente da Lincoln Paterson, della sede australiana della Commonwealth Scientific and Research Organization. Paterson fa notare che, formalmente, i fenomeni alla base dell'aggregazione per diffusione e della ramificazione viscosa sono identici. Nel primo caso la crescita è dovuta al flusso di particelle in movimento casuale diretto verso la struttura in accrescimento. Il verso del flusso è determinato dal fatto che le particelle hanno maggiore probabilità di muoversi dalle regioni più popolate (lontane) verso quelle meno popolate (in prossimità dell'aggregato), di quanta ne abbiano per muoversi in verso opposto. Il flusso risulta proporzionale alla velocità di variazione della densità di popolazione fuori dell'aggregato, fatto descritto dalla nota legge di Fick per la diffusione.

Nel caso della ramificazione digitiforme viscosa, la pressione all'interno della glicerina è analoga alla popolazione delle particelle. La pressione è massima all'interfaccia fra la bolla d'aria e la glicerina. Ciò produce un flusso di glicerina nel verso opposto a quello di crescita della bolla d'aria. L'entità del flusso è proporzionale alla velocità di cambiamento della pressione fuori della bolla. Le dita crescono perché il liquido si allontana più facilmente dalla loro punta. Si verifica quindi una instabilità di crescita simile a quella che si osserva nell'aggregazione per diffusione.

Un altro fenomeno è stato spiegato praticamente nello stesso modo. Quando si applica una elevata differenza di potenziale fra un elettrodo e un'emulsione fotografica o una polvere fine sparsa

sulla superficie di un isolante, si ottiene l'immagine della scarica elettrica dall'aspetto ramificato simile a un lampo (*si veda l'illustrazione in basso a sinistra a pagina 76*). Questa immagine è chiamata figura di Lichtenberg, in onore del fisico tedesco del XVIII secolo, Georg Christoph Lichtenberg. Nel 1984 un gruppo di ricercatori della Brown, Boveri & Company Limited, in Svizzera, ha segnalato che la crescita di una figura di Lichtenberg sembra dovuta a un fenomeno di aggregazione per diffusione.

Supponiamo che la tensione applicata inizialmente sia abbastanza elevata da danneggiare l'emulsione nei pressi dell'elettrodo, creando un canale di conduzione. Al di fuori del canale vi è un campo elettrico, la cui intensità è definita come la velocità di variazione della differenza di potenziale elettrico entro l'e-

La figura schematizza la crescita di un frattale mediante aggregazione per diffusione. Si inizia con un aggregato liscio, al quale aderiscono le particelle. Per il «rumore» dovuto al loro arrivo casuale, si formano delle piccole protuberanze e delle buche (*si veda l'illustrazione a sinistra*). Le linee nere mostrano il cammino delle particelle in arrivo, quelle in colore sono linee di densità media costante di particelle e quelle grige mostrano la direzione del loro flusso medio. Le protuberanze crescono più rapidamente delle buche (*in alto*) perché le particelle in arrivo hanno maggiore probabilità di aderire in vicinanza della loro sommità e, in ogni caso, quasi certamente aderiscono ai loro fianchi prima di raggiungere una buca. In questo modo le protuberanze crescono sempre più alte e il riempimento delle buche diviene sempre più improbabile. In questo modo, l'aggregato liscio si distorce rapidamente.

La simulazione al calcolatore dell'aggregazione per diffusione in uno spazio tridimensionale produce un frattale con dimensione 2,4 (*qui sopra*). Questa struttura somiglia strettamente a quella di un aggregato di rame della medesima dimensione (*a destra*). L'immagine al calcolatore è stata elaborata da Roy Richter dei General Motors Research Laboratories e l'aggregato di rame, che è stato ottenuto come precipitato da una soluzione di solfato rameico, è stato preparato da Nancy Hecker e David G. Grier dell'Università del Michigan.

mulsione. Il gruppo della Brown Boveri ha avanzato l'ipotesi che la crescita del canale sia più probabile nella direzione in cui il campo è più intenso; ciò accade proprio in corrispondenza delle punte più sottili della scarica. Quindi, le punte si allungano e proliferano secondo un accrescimento frattale.

L'elemento unificatore fra la deposizione di un metallo su un elettrodo, la ramificazione di un fluido viscoso e la formazione di una figura di Lichtenberg trova la sua migliore espressione nel linguaggio astratto delle equazioni differenziali alle derivate parziali. Ma non è difficile farsi un'idea dell'unitarietà del discorso ricorrendo a un'analogia con un foglio di gomma fissato lungo i quattro i lati e posto in tensione dal peso di un frattale che cresce a partire dal centro.

La probabilità di un certo passo nel cammino casuale di una particella, la pressione in una cella di Hele-Shaw e il potenziale elettrico presso un canale di conduzione entro un isolante sono tutti espressi da funzioni armoniche, che sono soluzioni di un sistema di equazioni differenziali. La superficie descritta da una funzione armonica ha sempre una curvatura totale nulla. Infatti, se essa si piega verso l'alto in una direzione, si piega contemporaneamente verso il basso nella direzione perpendicolare alla precedente, proprio come accade in una sella da cavallo. Anche un foglio di gomma deformato da un peso ha curvatura nulla. Quindi si può pensare che la superficie della buca prodotta dal frattale in accrescimento descriva l'andamento di una probabilità, una pressione o una tensione e che la sua pendenza in corrispondenza dei margini del frattale sia proporzionale alla sua velocità di crescita. Le pendenze maggiori si hanno proprio in corrispondenza delle protuberanze più appuntite, che quindi crescono maggiormente. Nello stadio di crescita successivo la gomma è premuta da protuberanze ancora più appuntite e così via.

È una tentazione quella di speculare su quanto lontano si possa spingere lo studio dei frattali. In effetti, le ramificazioni dei vasi sanguigni, delle vie aeree nei polmoni o dei vari tipi di corallo ricordano le strutture frattali che si formano mediante aggregazione per diffusione. Per quanto diversi ricercatori abbiano cercato di definire modelli di crescita per tali strutture, nessuno si è servito esplicitamente della geometria dei frattali. Se poi questa si dimostri utile nello studio dei fenomeni di crescita biologica è cosa ancora tutta da provare.

Il modello di aggregazione per diffusione è stato utilizzato per descrivere diversi altri fenomeni fisici, come la cristallizzazione superficiale di pellicole amorfe. Inoltre una generalizzazione del modello, detta condensazione di aggregati, ha consentito di descrivere la struttura di colloidi e di aerosol, come il nerofumo. In questo modello, proposto da Paul Meakin dalla E. I. du Pont de Nemours & Company Inc. e da Max Kolb, Rémi Jullien e Robert Botet dell'Università di Parigi a Orsay, si formano molti aggregati, che poi si muovono casualmente e si uniscono tra loro. In breve, i modelli di aggregazione si sono mostrati molto utili per descrivere vari sistemi fisici.

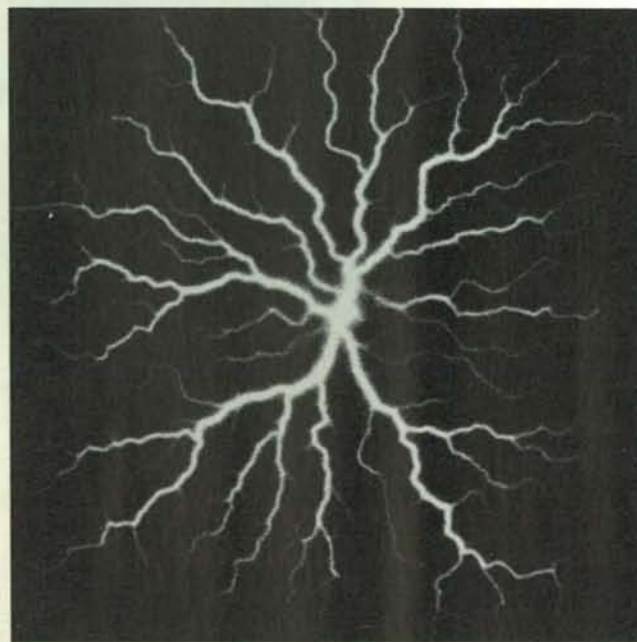
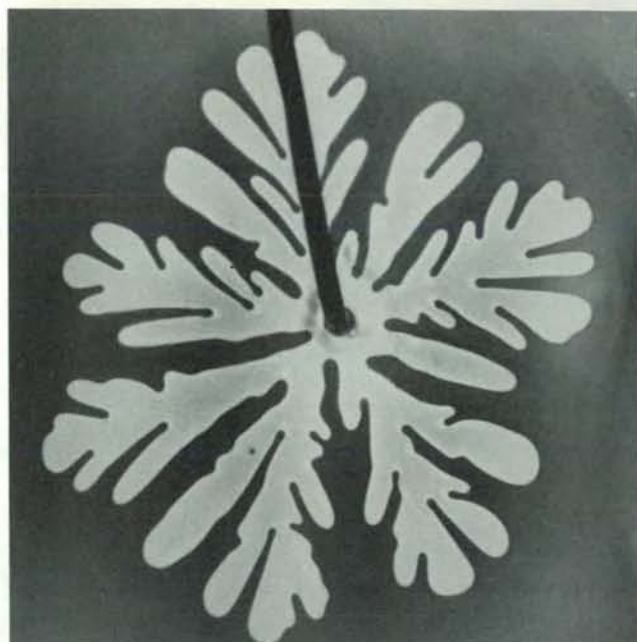
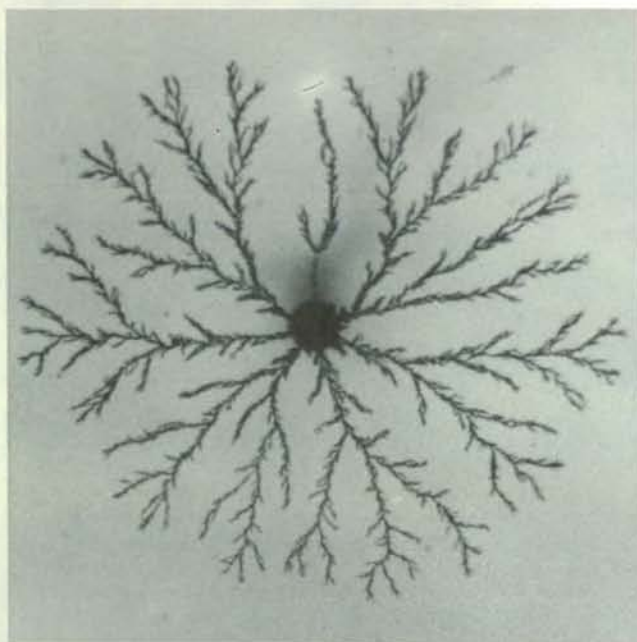
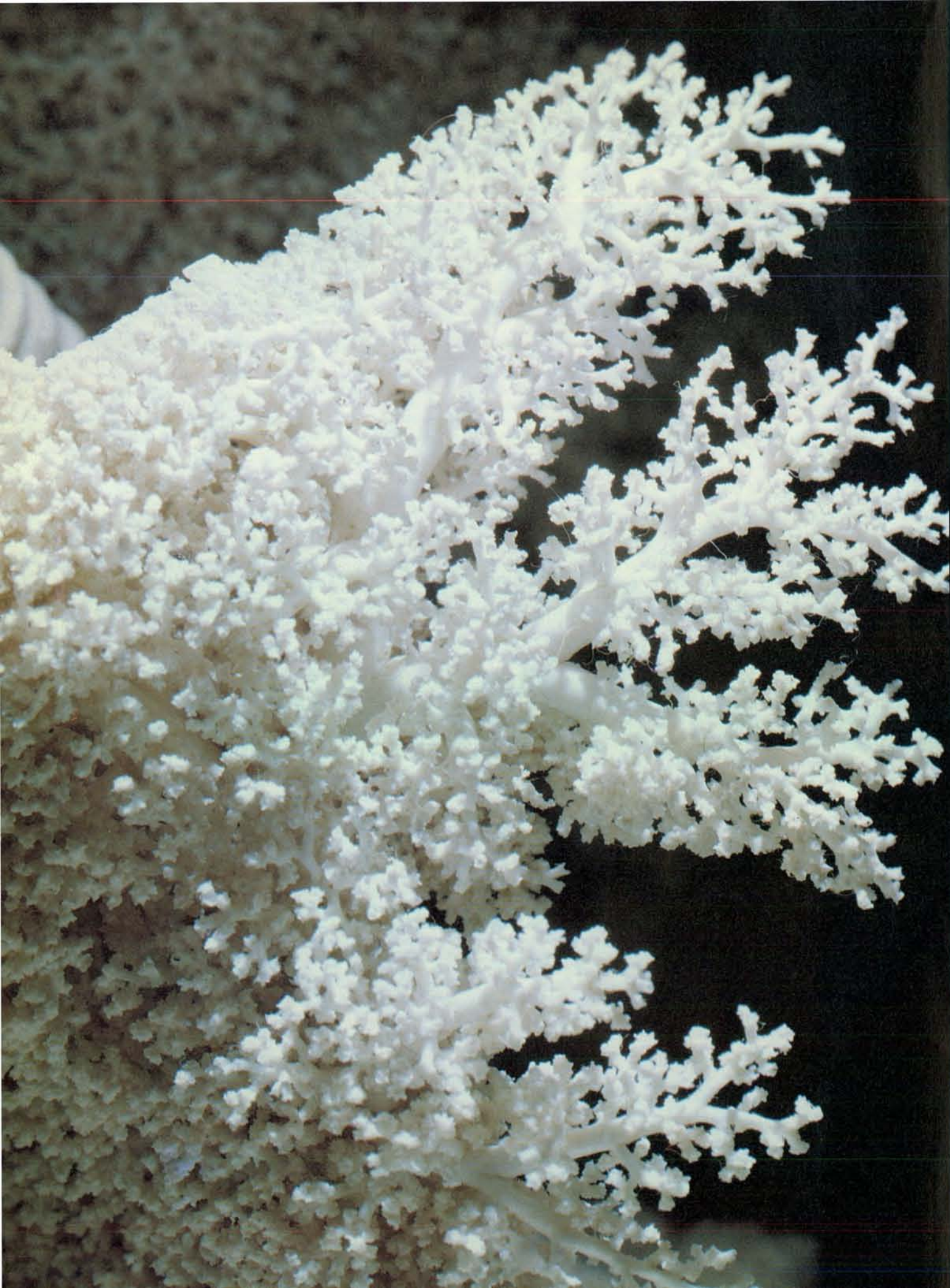
Ma allo stesso tempo va detto che i frattali non descrivono tutte le strutture ramificate che si trovano in natura. I cristalli di ghiaccio della neve, per esempio, probabilmente non sono dei frattali. Essi sono certamente strutture complesse, ma presentano una simmetria molto più evidente degli aggregati formati per diffusione e quindi appartengono a una categoria di cristalli detti dentriti. La struttura macroscopica di un qualunque cristallo di ghiaccio riflette l'anisotropia microscopica delle celle esagonali nelle quali si dispongono le molecole d'acqua. Ci si può allora chiedere come mai lo zinco, che pure cristallizza in un reticolo esagonale, si depositi con struttura frattale in una cella di elettrolisi (*si veda* l'illustrazione a pagina 76 in alto a sinistra). La risposta è che l'accrescimento, per quanto avvenga in condizioni di non equilibrio, è talmente lento che lo sdoppiamento delle punte cancella l'anisotropia del reticolo. Infatti si osserva che, quando sale la velocità di accrescimento in seguito a un aumento della differenza di potenziale nella cella, l'anisotropia si fa sentire e si ottiene quindi una struttura dendritica (*si veda l'illustrazione a pagina 76 in basso a destra*). Attualmente, diversi gruppi di ricercatori si stanno dedicando allo studio delle condizioni di transizione dalla struttura frattale a quella dendritica.

Fino a questo punto è stato preso in considerazione solo un particolare processo di crescita che produce strutture frattali: l'aggregazione per diffusione. È possibile mettere a frutto le conoscenze acquisite? In particolare, le proprietà di scala fin qui esposte conducono alla comprensione di proprietà fisiche degli aggregati diverse da quelle puramente geometriche?

Cominciano a emergere degli indizi che fanno ben sperare. Per esempio, negli ultimi anni Raoul Kopelman e collaboratori hanno studiato all'Università del Michigan diverse reazioni chimiche che si osservano negli aggregati di percolazione, i quali sono frattali che si formano all'equilibrio. Questi ricercatori hanno dimostrato che la reazione ha un comportamento singolare quando è confinata nell'aggregato. A differenza di

Un foglio di gomma, stirato e fissato lungo i lati e teso dal peso di un frattale in accrescimento, costituisce un semplice modello meccanico dell'aggregazione per diffusione. Il frattale cresce più in fretta dove le pareti della buca sono più ripide, cioè in corrispondenza delle sue protuberanze. Nella fase successiva il foglio sarà teso da protuberanze ancora più sporgenti e così via.

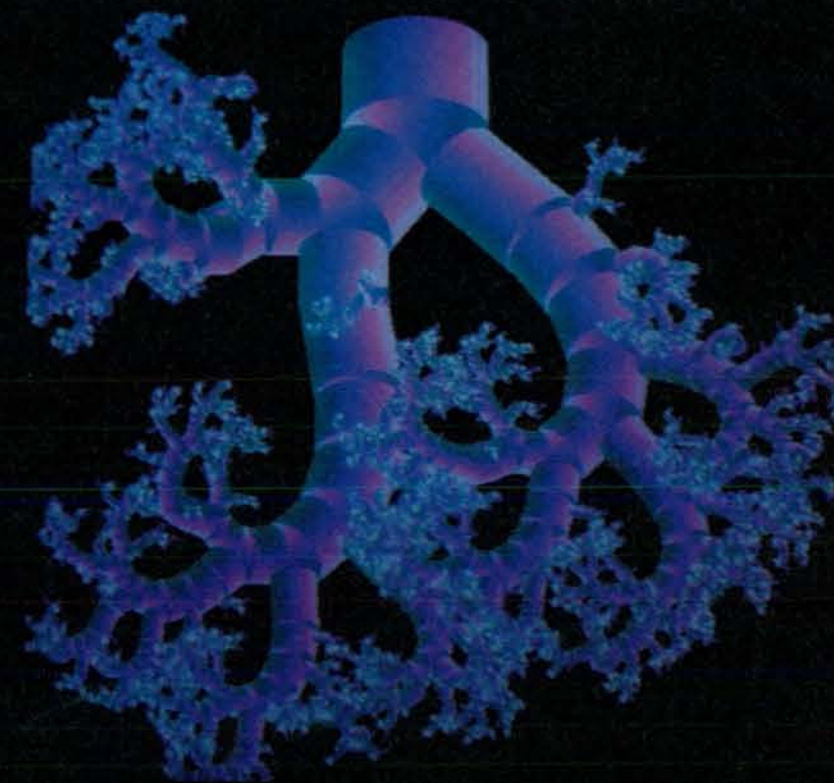
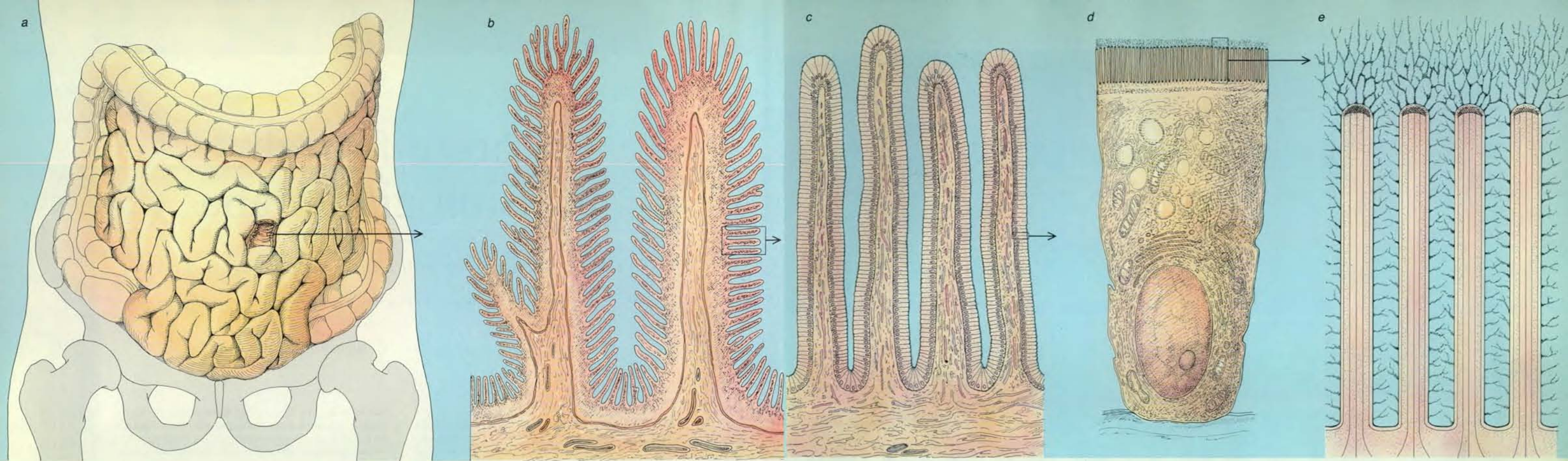
una reazione chimica tipica, che procede con velocità costante, la velocità di reazione in un aggregato di percolazione varia nel tempo. Il motivo fondamentale deve essere ricercato nel fatto che le specie chimiche che diffondono lungo un frattale non si muovono altrettanto facilmente di come farebbero in fase libera. L'incontro fra le specie che devono reagire è reso meno probabile dal fatto che devono migrare lungo una struttura con molti vicoli ciechi.

La velocità delle reazioni che avvengono entro un frattale dipende sia dalla dimensione frattale, sia dal modo in cui i reagenti si muovono lungo l'aggregato. I due fattori insieme danno origine a un altro parametro caratteristico, denominato dimensione spettrale. Questo parametro è stato introdotto da Shlomo Alexander della Hebrew University di Gerusalemme e da Raymond L. Orbach dell'Università della California a Los Angeles, con lo scopo di descrivere la diffusione e la dinamica su un frattale. Per quanto attualmente non siano disponibili dati sperimentali riguardanti la dimensione spettrale nei frattali cresciuti lontano dall'equilibrio, vi sono tutte le ragioni per credere che un siffatto parametro esista. È probabile che nel prossimo futuro il semplice fatto che esista la geometria dei frattali consenta di aprire parecchie frontiere della fisica che al giorno d'oggi vengono ancora considerate invalicabili.

In natura i frattali paiono crescere per diffusione e aggregazione. L'illustrazione mostra un deposito elettrolitico di zinco (*in alto a sinistra*), la «ramificazione digitiforme viscosa» di una bolla d'aria nella glicerina (*in alto a destra*) e la struttura di una scarica elettrica nota come figura di Lichtenberg (*in basso a sinistra*). La striscia terminante al centro della bolla è il tubo di adduzione dell'aria. L'aggregato di zinco in basso a destra mostra che cosa succede aumentando la tensione applicata alla cella elettrolitica: la struttura diventa dendritica, simile a un cristallo di neve. I depositi di zinco sono stati ottenuti da D. G. Grier e la ramificazione viscosa da E. Ben-Jacob dell'Università del Michigan. La figura di Lichtenberg è di L. Niemeyer e H. J. Wiesmann della Brown, Boveri & Co., Ltd., e di L. Pietronero dell'Università di Groningen.

Caos e frattali in fisiologia umana

Il caos nella funzione biologica è proprio dello stato di salute, mentre la periodicità può essere patologica

di Ary L. Goldberger, David R. Rigney e Bruce J. West

Un monitoraggio del ritmo cardiaco mostra come la frequenza dei battiti cambi a volte in modo sensibile da un minuto all'altro o da un'ora all'altra; una broncoscopia consente di osservare come la trachea si ramifichi in vie aeree sempre più strette. Di fronte a questi fenomeni si può arrivare a percepire come l'intervallo tra i battiti cardiaci vari in modo caotico e la ramificazione delle vie aeree assomigli a un frattale. Solo recentemente, però, fisiologi e clinici hanno cominciato a quantificare la possibilità dell'esistenza di dinamiche caotiche e architetture frattali; nuove ricerche stanno mettendo in discussione principi dati per scontati in medicina e si stanno riconoscendo nuovi potenziali segni premonitori di malattia. L'interpretazione convenzionale in medicina vuole che la malattia e la vecchiaia siano l'esito del logoramento di un sistema ordinatamente deterministico e che questo logoramento riduca l'ordine, provocando risposte erratiche dell'organismo o scom-

Le vie aeree polmonari (*a sinistra*), così come sono state formate dall'evoluzione e dallo sviluppo embrionale, ricordano frattali generati al calcolatore (*in basso*). Bronchi e bronchioli (che qui sono mostrati in un calco) formano un albero con ramificazioni multiple, la cui configurazione si presenta simile sia ad alto che a basso ingrandimento. Misurando i diametri dei diversi ordini di ramificazione, si è appurato che l'albero bronchiale può essere descritto con la geometria frattale.

bussolandone i normali ritmi periodici.

Negli ultimi cinque anni, noi e altri nostri colleghi abbiamo scoperto che il cuore e altri sistemi fisiologici possono comportarsi erraticamente, specialmente quando sono giovani e in buona salute. Al contrario, un aumento della regolarità del funzionamento si accompagna a volte alla vecchiaia e alla malattia.

Irregolarità e imprevedibilità sono caratteristiche importanti dello stato di salute; mentre una variabilità diminuita e periodicità accentuate sono associate a molte malattie.

Stimolati da queste ipotesi, noi e altri fisiologi abbiamo cercato periodicità che potessero segnalare l'evoluzione di stati patologici (specialmente malattie cardiache). Inoltre abbiamo cominciato ad analizzare la flessibilità e la consistenza di strutture frattali irregolari e l'adattabilità e stabilità di sistemi che si comportano in modo apparentemente caotico.

Caos e frattali sono argomenti associati allo studio dei fenomeni non lineari, ossia dei sistemi che rispondono sproporzionatamente a stimoli. La teoria della dinamica non lineare ha fornito chiarimenti sui fenomeni epidemici, sulla cinetica di alcune reazioni chimiche e sui mutamenti del tempo atmosferico. In certe condizioni, sistemi non lineari deterministici - quelli costituiti soltanto da pochi semplici elementi [nessuno dei quali aleatorio, *ndt*] - si comportano in modo erratico: uno stato chiamato caos. Il caos deterministico dei sistemi dinamici non lineari non è l'analogo del caos nel senso letterale di completa disorganizzazione e casualità. Il caos non lineare si riferisce a un tipo di casualità vincolata che può essere associata con la geometria frattale.

Le strutture frattali sono spesso esito di dinamiche non lineari caotiche: quando fenomeni caotici modellano l'ambiente (la risacca del mare, le turbolenze atmosferiche, i movimenti geologici), ne rimangono a testimonianza forme frattali (linee costiere, nubi, formazioni rocciose). Ciononostante la matematica dei frattali si è sviluppata indipendentemente da quella delle dinamiche non lineari e anche oggi le connessioni fra le due discipline non sono del tutto definite.

Un frattale consiste di frammenti geometrici di grandezza e di orientamento variabile, ma con forma simile. Certi neuroni, per esempio, hanno una struttura simile ai frattali. Se si esaminano tali neuroni a basso ingrandimento si possono osservare ramificazioni asimmetriche, i dendriti, connesse con i corpi cellulari. A ingrandimento leggermente superiore si osservano ramificazioni più piccole a partire da quelle più grandi. Aumentando ulteriormente l'ingrandimento, si può scorgere un altro livello di dettaglio: ramificazioni di ramificazioni di ramificazioni. Anche se la ramificazione del neurone a un certo livello si ferma, i frattali ideali possiedono infiniti dettagli (si veda l'articolo *L'accrescimento dei frattali* di Leonard M. Sander in «Le Scienze» n. 223, marzo 1987). Forse ancora più notevole è il fatto che i dettagli di un frattale, a una certa scala, sono simili (seppur non necessariamente identici) a quelli della struttura vista a ingrandimento maggiore o minore.

Di fronte a due fotografie di dendriti a differenti livelli di ingrandimento (senza alcun riferimento di scala sull'immagine), può essere difficile riconoscere a quale ingrandimento corrisponda una determinata fotografia. Questa proprietà interna, comune a tutti i frattali, è chiamata autosomiglianza.

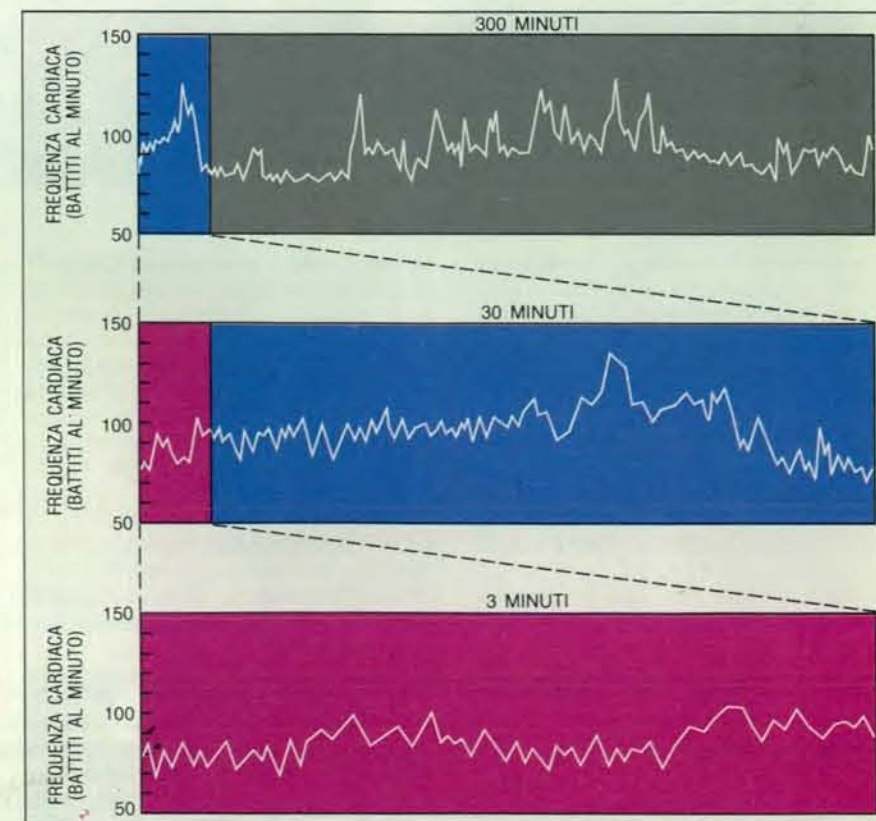
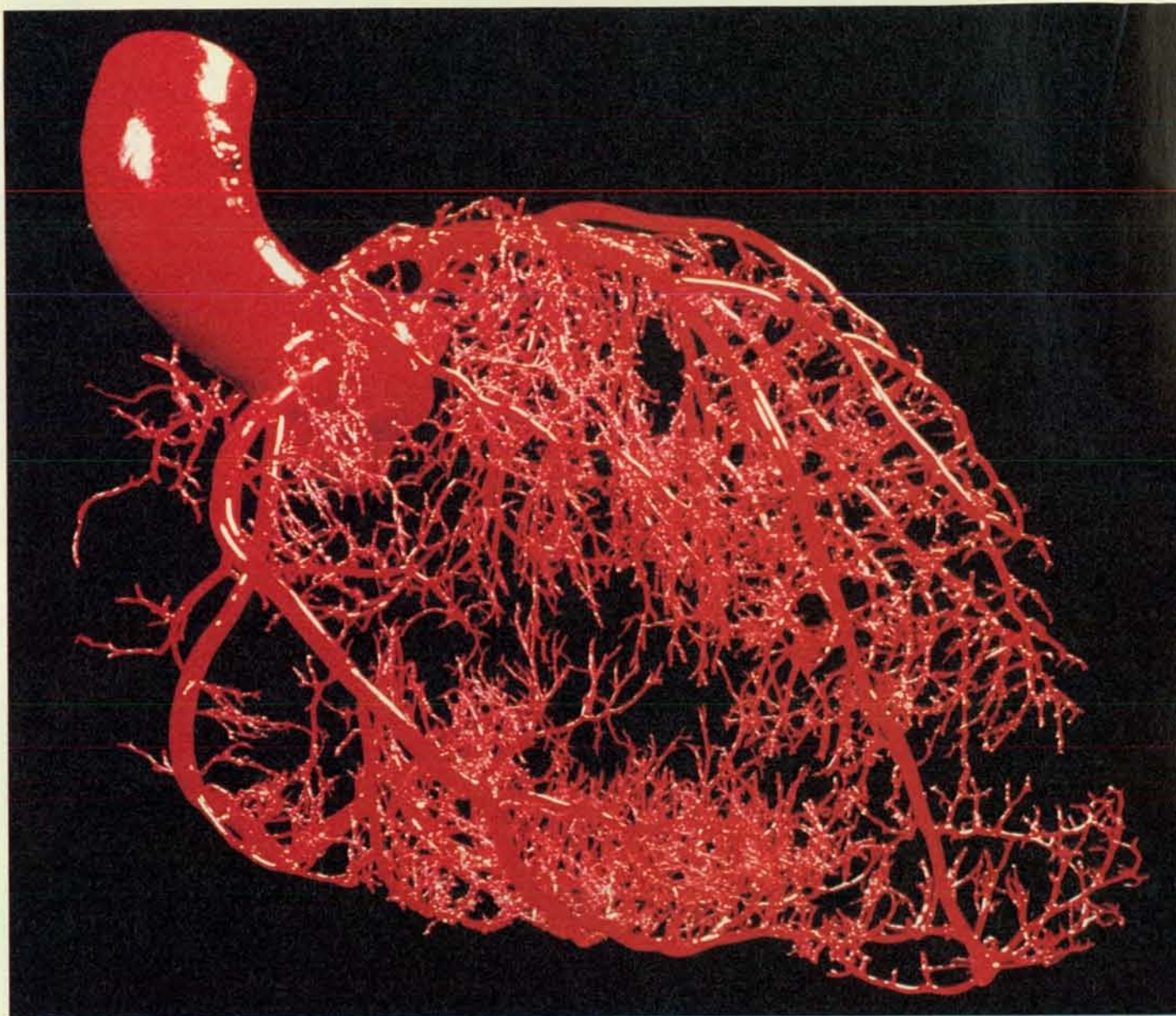
L'essere composto da dettagli autosimili a qualsiasi ingrandimento fa sì che il frattale non abbia lunghezza definita. Se si prova a misurare la lunghezza di un frattale con un righello, costruito in base a una data unità di misura, alcuni dettagli saranno comunque più piccoli di quanto l'unità di misura possa misurare. Pertanto, al crescere della risoluzione la lunghezza di un frattale aumenta.

Dato che la lunghezza non è per i frattali un concetto significativo, i matematici calcolano la «dimensione frattale» per quantificare quanto spazio venga occupato da essi. Il concetto di dimensione si applica convenzionalmente agli oggetti della geometria classica, o euclidea. Le linee hanno dimensione uno, i cerchi hanno dimensione due, le sfere hanno dimensione tre. Ma i frattali hanno dimensioni [anche, *ndt*] frazionarie. Mentre una linea regolare euclidea occupa precisamente uno spazio a una dimensione, la linea frattale si dipana all'interno di uno spazio a due dimensioni. Una linea frattale - una linea costiera, per esempio - ha perciò dimensione compresa fra uno e due. Allo stesso modo una superficie frattale - una montagna, per esempio - ha una dimensione che varia fra due e tre. Maggiore è la dimensione di un frattale, maggiore è la probabilità che una determinata regione di spazio contenga una porzione di quel frattale.

Nel corpo umano strutture riconducibili ai frattali sono osservabili nelle reti di vasi sanguigni, di fibre nervose e di strutture canalizzate. Il frattale più attentamente studiato è l'albero bronchiale, il sistema ramificato che trasporta i gas respiratori da e verso i polmoni. Nel 1962, Ewald R. Weibel e Domingo M. Gomez e più tardi Otto G. Raabe e i suoi collaboratori eseguirono dettagliate misurazioni dei diametri interni delle vie aeree. Recentemente due di noi (West e Goldberger) in collaborazione con Valmik Bhargava e Thomas R. Nelson dell'Università della California a San Diego hanno rianalizzato quelle misurazioni, effettuate su calchi di polmone di uomo e di altre specie di mammiferi.

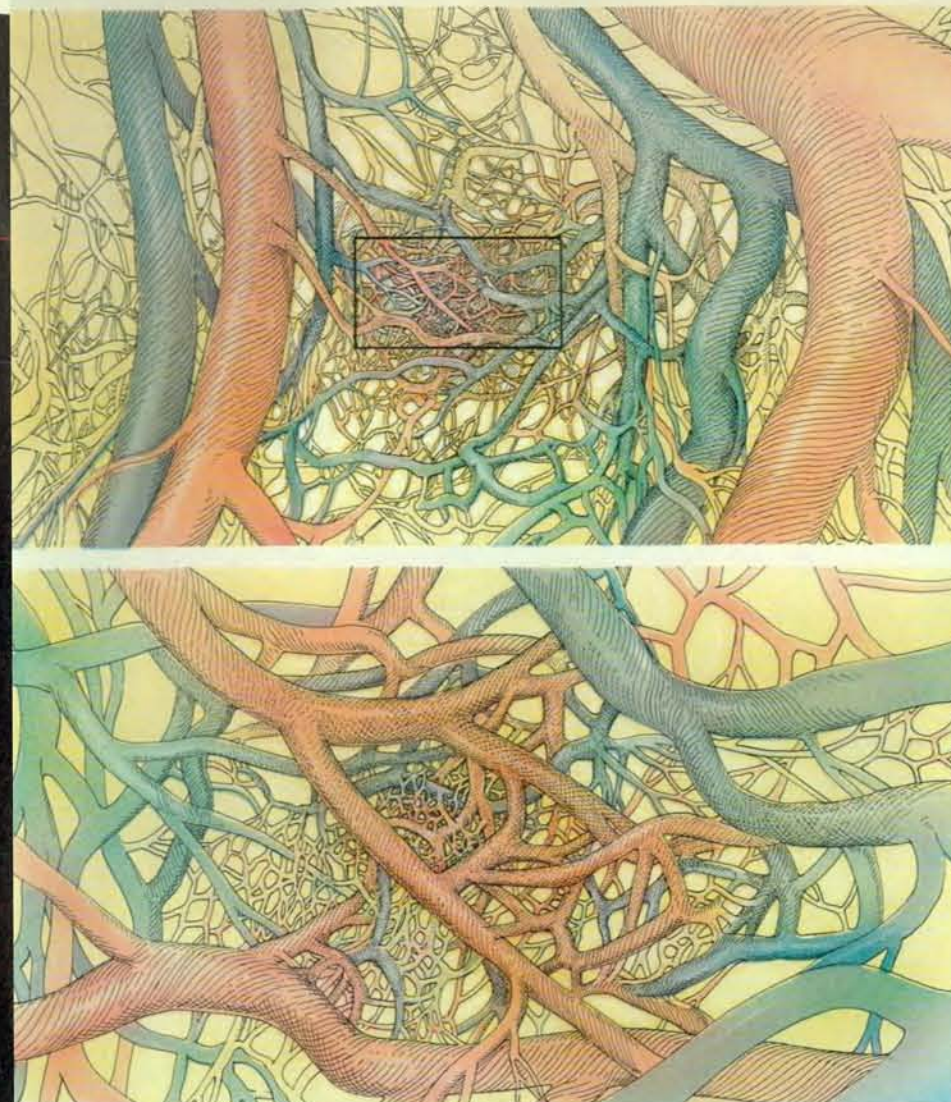
Pur con piccole differenze tra le diverse specie, i risultati mostrano tra queste misurazioni rapporti tipici di oggetti frattali. Anche molti altri organi sembrano frattali, sebbene le loro dimensioni non siano state ancora quantificate.

Nel cuore, le strutture frattali hanno un ruolo vitale nella meccanica della contrazione e nella conduzione dello stimolo elettrico eccitatorio. Per esempio una rete frattale di arterie e vene coronarie trasporta sangue da e verso il cuore. James B. Bassingthwaighte e Hans

Autosomiglianza in un sistema vuol dire che le caratteristiche di quel sistema, sia esso una struttura o un processo, si presentano simili a ogni scala di lunghezza o di tempo. Un'osservazione dell'intestino tenue a ingrandimenti diversi (*in alto*) fa pensare appunto all'autosomiglianza. Quando la frequenza cardiaca in un individuo sano viene registrata per 3, 30 e 300 minuti (*curve in basso*), i relativi grafici mostrano lo stesso tipo di fluttuazioni.

van Beek dell'Università di Washington hanno recentemente utilizzato la geometria frattale per spiegare alcune anomalie nelle modalità del flusso sanguigno coronarico, la cui interruzione può causare l'infarto miocardico. Inoltre un intreccio frattale di fibre di tessuto connettivo all'interno del cuore - le corde tendinee - lega le valvole mitrale e tricuspide al muscolo sottostante. Se questi tessuti dovessero rompersi, vi sarebbe un forte rigurgito di sangue dai ventricoli agli atrii, seguito da insufficienza cardiaca congestizia. Infine l'architettura frattale è evidente anche nelle ramificazioni di certi muscoli cardiaci così come nel sistema di Purkinje, che conduce i segnali elettrici dagli atrii ai muscoli cardiaci dei ventricoli.

Nonostante che queste strutture anatomiche di tipo frattale facciano da supporto a funzioni evidentemente diverse in organi differenti, esse hanno in comune tratti strutturali e fisiologici. Ramificazioni o ripiegature frattali amplificano grandemente la superficie disponibile per l'assorbimento (come nel caso dell'intestino), la distribuzione e la raccolta dei soluti (vasi sanguigni, dotti biliari e albero bronchiale) e per l'elaborazione delle informazioni (reti di neuroni). Le strutture frattali, anche in virtù della loro ridondanza e irregolarità, sono piuttosto resistenti a eventuali danni. Il cuore, per esempio, può continuare a pompare con uno svantaggio meccanico relativamente modesto anche in caso di un danno esteso al sistema di Purkinje.

Le strutture frattali del corpo umano si originano attraverso una dinamica lenta di sviluppo ed evoluzione embrionale. Abbiamo ipotizzato che tali processi - come altri che producono strutture frattali - mostrino un caos deterministico. Recenti ricerche in campo fisiologico hanno portato alla luce altri esempi di dinamiche apparentemente caotiche in scale temporali di più breve durata, sperimentalmente «accessibili».

All'inizio degli anni ottanta, quando i ricercatori cominciarono ad applicare la teoria del caos ai sistemi fisiologici, ci si aspettava che il caos sarebbe stato maggiormente osservabile negli stati patologici e nella vecchiaia. In definitiva, l'intuizione e la tradizione medica avevano sostenuto fino ad allora una visione di questo tipo. Se si ausculta il cuore attraverso un fonendoscopio o si rileva il polso, il ritmo cardiaco sembra essere regolare. Per un individuo a riposo l'intensità e la frequenza delle pulsazioni sembrano abbastanza costanti. Per questo motivo i cardiologi descrivono ordinariamente la frequenza cardiaca normale come «ritmo sinusale normale». Un'analisi più attenta rivela che individui sani hanno frequenze cardiache che variano considerevolmente anche a riposo. In giovani adulti sani, la frequenza cardiaca, che in media è di circa 60 battiti al minuto, può variare addirittura di 20 battiti al minuto nel giro di pochi secondi. In una giornata la frequenza cardiaca può passare da quaranta a centottanta battiti al minuto.

Per almeno mezzo secolo i medici hanno interpretato le fluttuazioni di frequenza cardiaca in termini di omeostasi: i sistemi fisiologici normalmente si comportano in modo da ridurre la variabilità e mantenere costanti le funzioni interne.

Secondo questa teoria sviluppata da Walter B. Cannon della Harvard Medical School, qualsiasi variabile fisiologica, inclusa la frequenza cardiaca, dovrebbe ritornare al suo stato stazionario dopo essere stata perturbata. Il principio dell'omeostasi suggerisce che le variazioni della frequenza cardiaca sono semplicemente risposte transitorie a un ambiente oscillante. Si potrebbe ragionevolmente immaginare che, nella malattia o nella vecchiaia, l'organismo perda la capacità di mantenere una frequenza cardiaca costante a riposo e che quindi la variabilità aumenti.

Ma le cose si presentano altrimenti a chi misuri scrupolosamente i normali intervalli tra battiti cardiaci e li registri in tutto l'arco di una giornata. La serie temporale ottenuta sembra irregolare e, a prima vista, completamente casuale. Esaminando però il grafico a differenti scale temporali, emergono alcune caratteristiche. Se ci si concentra sulla serie temporale relativa all'arco di poche ore, si riscontrano fluttuazioni più rapide, le cui escursioni e la cui sequenza sembrano in qualche modo simili a quelle trovate in serie temporali più lunghe. A scale temporali ancora più ristrette (minuti), è possibile osservare fluttuazioni ancora più rapide che sembrano sempre molto simili all'andamento iniziale.

I vasi sanguigni del cuore (qui mostrati in un calco) presentano ramificazioni di tipo frattale. I vasi principali visibili a sinistra si ramificano in una serie di vasi più piccoli (*qui sopra in alto*) che, a loro volta, si ramificano in vasi di calibro ancora più ridotto.

Le fluttuazioni tra battiti contigui in scale temporali differenti appaiono autosimili, esattamente come le ramificazioni di un frattale geometrico. Questo risultato suggerisce che i meccanismi di controllo della frequenza cardiaca possano essere intrinsecamente caotici. In altre parole, la frequenza può oscillare considerevolmente, anche in assenza di stimoli esterni oscillanti, anziché portarsi a un omeostatico stato stazionario.

Un metodo per stabilire se le variazioni nella frequenza cardiaca siano caotiche o periodiche è quello di calcolare lo spettro di Fourier della serie temporale. Lo spettro di Fourier di una forma d'onda qualsiasi (come una serie temporale) rivela la presenza di componenti periodiche. Se una serie temporale mostra una frequenza cardiaca di un battito esatto al secondo, lo spettro relativo sarebbe costituito da un unico picco situato alla frequenza di un battito al secondo. D'altra parte, una serie temporale di battiti cardiaci caotici dovrebbe generare uno spettro costituito sia da picchi larghi, sia da altri picchi non meglio connotati. In realtà l'analisi spettrale della frequenza cardiaca normale mostra un ampio spettro che ricorda una situazione caotica. Un altro strumento per analizzare le dinamiche di sistemi complessi non lineari è la rappresentazione nello spazio delle fasi o spazio degli stati. Questa tecnica segue i valori delle variabili indipendenti che cambiano nel tempo. Il numero e il tipo delle variabili indipendenti è tipico di ogni singolo sistema (si veda l'articolo *Il caos* di James P. Crutchfield, J. Doyne Farmer, Norman H. Packard e Robert S. Shaw in «Le Scienze» n. 222, febbraio 1987).

Il gran numero di variabili indipendenti presenti in molti sistemi complessi le rende non immediatamente identificabili e misurabili. Per tali sistemi la rappresentazione nello spazio delle fasi può essere realizzata usando il metodo delle *delay maps* [dette anche *return maps*, *ndt*]. Nella più semplice *delay map* l'ascissa di ogni punto corrisponde al valore di una variabile in un certo istante mentre l'ordinata crea il valore della stessa variabile dopo un ritardo costante. Una serie di questi punti in tempi successivi delinea una curva o traiettoria che descrive l'evoluzione del sistema.

Per identificare il tipo di sistema dinamico, caotico o periodico, si determinano le traiettorie per differenti condizioni iniziali. Successivamente si cerca un attrattore, una regione dello spazio delle fasi che attrae le traiettorie.

Il tipo più semplice di attrattore è il punto fisso. Esso descrive un sistema - come un pendolo smorzato - che evolve sempre verso un singolo stato. Nello spazio delle fasi vicino all'attrattore a punto fisso tutte le traiettorie convergono verso quel singolo punto.

Una forma più complicata di attrattore è il ciclo limite. Esso corrisponde a un

sistema - come un pendolo ideale senza attriti - che evolve verso uno stato periodico. Nello spazio delle fasi vicino al ciclo limite le traiettorie seguono un percorso regolare che può essere circolare o ellittico.

Altri attrattori, definiti semplicemente «attrattori strani», descrivono sistemi che non sono né statici né periodici. Nello spazio delle fasi vicino a un attrattore strano, due traiettorie che presentano condizioni iniziali praticamente identiche divergono rapidamente e su lunghe distanze temporali divengono molto differenti. Un sistema la cui descrizione presenta attrattori strani è caotico.

Abbiamo recentemente analizzato la rappresentazione dello spazio delle fasi per il battito cardiaco normale. I risultati mostrano un comportamento più simile a un attrattore strano che non a un attrattore periodico, caratteristico di un processo realmente regolare. Queste osservazioni concordano con le indagini cliniche che hanno dimostrato come la dinamica di un battito cardiaco normale possa essere caotica.

Il meccanismo responsabile di una variabilità caotica nella frequenza cardiaca dell'individuo sano nasce probabilmente entro il sistema nervoso. Il nodo senoatriale (il pacemaker naturale del cuore) riceve il segnale dalla parte involontaria del sistema nervoso, il sistema nervoso autonomo; questo, a sua volta, ha due componenti principali: il sistema parasimpatico e quello simpatico. La stimolazione parasimpatica diminuisce la frequenza di scarica delle cellule del nodo senoatriale, mentre una stimolazione simpatica ha l'effetto opposto. Queste due componenti agiscono come un continuo tiro alla fune sul pacemaker. Le fluttuazioni della frequenza cardiaca nei soggetti sani sono il risultato di questo continuo alternarsi di stimoli.

Recentemente alcuni ricercatori, tra i quali Richard J. Cohen e colleghi del Massachusetts Institute of Technology, hanno quantificato la riduzione della variabilità del battito cardiaco a seguito di un intervento di trapianto cardiaco nel corso del quale le fibre nervose del sistema autonomo vengono recise.

La frequenza cardiaca viene qui mostrata in grafici delle serie temporali (*a sinistra*), in spettri di Fourier (*al centro*) e in rappresentazioni nello spazio delle fasi (*a destra*). Una frequenza registrata 13 ore prima di un arresto (*in alto*) è quasi costante: lo spettro è piatto e la traiettoria nello spazio delle fasi fa pensare a un attrattore a punto fisso. Una frequenza che precede di otto giorni una morte cardiaca improvvisa (*al centro*) è abbastanza periodica: nello spettro vi è un solo picco e nello spazio delle fasi si ha traiettoria a ciclo limite. La frequenza di un cuore sano (*in basso*) appare erratica: ha spettro allargato e una traiettoria del tipo ad attrattore strano.

Osservazioni compiute in diversi laboratori suggeriscono che il caos è una caratteristica normale di altre componenti del sistema nervoso. Gottfried Mayer-Kress al Los Alamos National Laboratory, Paul E. Rapp del Medical College of Pennsylvania e Agnes Babloyantz e Alain Destexhe dell'Università libera di Bruxelles hanno analizzato gli elettroencefalogrammi di giovani soggetti sani e hanno trovato chiare prove di caos nel sistema nervoso.

Otto E. Rössler e i suoi colleghi dell'Università di Tübingen hanno anche scoperto segni di caos in componenti del sistema nervoso responsabili della secrezione ormonale. Essi hanno analizzato le variazioni temporali del livello ormonale in soggetti sani e hanno trovato situazioni a quanto pare caotiche.

Altri ricercatori hanno simulato l'interazione tra cellule nervose, per mostrare come possa esordire il caos. Walter J. Freeman all'Università di California a Berkeley ha dimostrato come il caos possa instaurarsi in un modello del sistema olfattivo. Il modello comprende un anello a retroazione fra i neuroni e un ritardo nei tempi di risposta. Precedentemente, Leon Glass e Michael C. Mackey della McGill University avevano riconosciuto l'importanza dei ritardi temporali per la produzione di caos.

Per quale motivo la frequenza cardiaca e altri sistemi controllati dal sistema nervoso dovrebbero avere dinamiche caotiche? Probabilmente perché tali dinamiche offrono vantaggi funzionali. I sistemi caotici agiscono in un ampio spettro di condizioni e pertanto sono adattabili e flessibili; queste caratteristiche consentono ai sistemi di rispondere alle necessità poste da un ambiente imprevedibile e in continua modificazione.

Molte patologie mostrano un aumento di periodicità e una diminuzione di variabilità. Le prime indicazioni del fatto che perfino un cuore sul punto di fermarsi può comportarsi con periodicità vengono dall'analisi di Fourier su forme d'onda di elettrocardiogrammi durante tachicardia parossistica o fibrillazione ventricolare, ritmi molto rapidi che assai frequentemente portano all'arresto cardiaco. Verso la metà degli anni ottanta, Raymond E. Ideker e i suoi colleghi alla Duke University School of Medicine registrarono le forme d'onda associate a fibrillazione ventricolare negli strati più interni del cuore di un cane. Essi riscontrarono che l'attività fibrillatoria all'interno del cuore è un fenomeno molto più periodico di quanto si ritenesse.

Nel 1988, due di noi (Goldberger e Rigney) hanno condotto uno studio retrospettivo sugli elettrocardiogrammi di persone che avevano avuto gravi patologie cardiache. Si scoprì che la frequenza cardiaca di tali pazienti spesso diventava meno variabile del normale in un momento qualsiasi, da minuti a mesi, prima della morte per arresto cardiaco. In certi casi la variabilità globale della frequenza cardiaca era ridotta; in altri comparivano oscillazioni periodiche della frequenza che poi si interrompevano improvvisamente. In modo quasi identico, il sistema nervoso può mostrare perdita di variabilità e insorgenza di periodicità patologiche in disordini come l'epilessia, il morbo di Parkinson e la sindrome maniaco-depressiva. Inoltre, mentre in soggetti sani il conteggio dei globuli bianchi è stato visto variare caoticamente da un giorno all'altro, si è osservato che in certi casi di leucemia il numero di globuli bianchi oscilla con periodicità.

L'andamento periodico in relazione alla malattia e quello apparentemente caotico in relazione alla salute non implicano che tutte le patologie siano associate a un aumento della regolarità. In certe aritmie cardiache il polso è così variabile che lo stesso paziente si lamenta delle cosiddette palpitazioni. Alcuni fra questi eventi rappresentano oscillazioni che sembrano irregolari, ma a un esame più attento si rivelano periodiche.

In altre aritmie il battito cardiaco è in effetti imprevedibile. Tuttavia nessuna di queste patologie irregolari si è dimostrata un esempio di caos non lineare, per quanto le pulsazioni possano sembrare «caotiche» in senso colloquiale.

In conclusione, la fisiologia sembra uno dei settori più promettenti per lo studio dei frattali e del caos, e di altri tipi di dinamiche non lineari. I fisiologi hanno bisogno di sviluppare una migliore comprensione di come i processi di sviluppo possano portare alla formazione di architetture frattali e di come i processi dinamici che si svolgono all'interno del corpo umano generino caos apparente.

Nel prossimo futuro gli studi sui frattali e sul caos applicati alla fisiologia potranno fornire metodi più sensibili per caratterizzare le disfunzioni dovute a vecchiaia, malattia o sostanze tossiche.

I neuroni sono un esempio di struttura frattale. Il corpo cellulare si ramifica in dendriti che si ramificano a loro volta. Questa struttura può essere correlata al caos nel sistema nervoso.

BIBLIOGRAFIA

WEST B. J., *An Essay on the Importance of Being Nonlinear* in *Lecture Notes in Biomathematics 62*, a cura di S. Levine, Springer-Verlag, 1985.

GOLDBERGER ARY L. e WEST BRUCE J., *Fractals in Physiology and Medicine* in «Yale Journal of Biology and Medicine», 60, pp. 421-435, 1987.

WEST BRUCE J. e GOLDBERGER ARY L., *Physiology in Fractal Dimensions* in «American Scientist», 75, n. 4, luglio-agosto 1987.

GOLDBERGER A. L., RIGNEY D. R., MIETUS J., ANTMAN E. M. e GREENWALD S., *Nonlinear Dynamics in Sudden Cardiac Death Syndrome: Heartrate Oscillations and Bifurcations* in «Experientia», 44, pp. 983-987, 1988.

MINETTI ALBERTO E., *I frattali in biomedicina* in *Scienza & Tecnica 89/90*, Arnoldo Mondadori, Milano, 1989.

I frattali e le linee suturali delle ammoniti

L'evoluzione delle linee suturali delle ammoniti verso forme frattali, vascolari e peaniane ha reso più efficiente la distribuzione degli sforzi sulla conchiglia aumentando la sua resistenza alla pressione idrostatica

di Giuseppe Damiani

A molte persone sarà capitato di osservare delle strane forme a spirale pietrificate in alcune stratificazioni di rocce sedimentarie o esposte nelle vetrine di un museo oppure nelle illustrazioni di qualche libro. Si tratta dei resti fossili delle conchiglie di molluschi marini della classe dei cefalopodi e più precisamente delle sottoclassi dei nautiloidi, di cui esistono ancora oggi rappresentanti viventi del genere *Nautilus*, e degli ammonoidi, chiamati più comunemente ammoniti, che sono i più diffusi e caratteristici macrofossili di invertebrati del Mesozoico e che si estinsero in massa nel Cretaceo, circa 60 milioni di anni or sono (si veda l'articolo *L'estinzione delle ammoniti* di Peter Ward in «Le Scienze» n. 184, dicembre 1983). Nautiloidi e ammoniti erano dotati di una conchiglia esterna costituita prevalentemente da carbonato di calcio, dalla forma di un cono senza base avvolto a spirale, che spesso si è conservata molto bene fossilizzandosi.

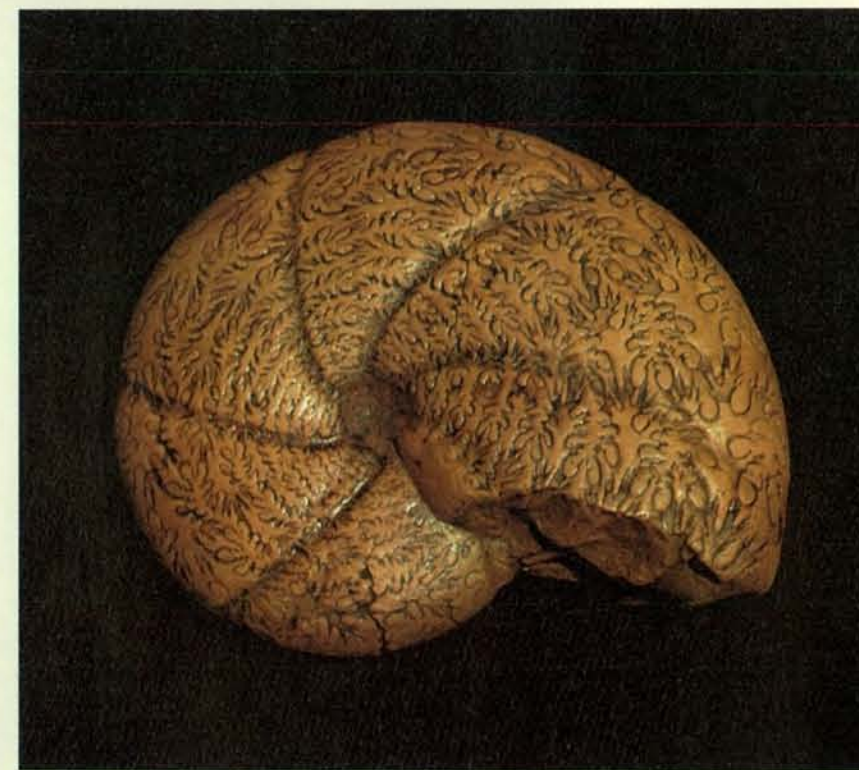
Solo pochi dei più curiosi e attenti osservatori avranno notato la presenza su queste conchiglie fossili di strane linee, che nelle ammoniti possono anche essere molto complesse. Queste linee vengono chiamate «linee suturali» o «lobali» e sono molto studiate dai paleontologi perché, data la loro complessità e specificità, sono di grande importanza per la classificazione e la determinazione delle relazioni filogenetiche dei nautiloidi e delle ammoniti. Le linee suturali sono visibili solo nei modelli interni fossili delle conchiglie, dato che sono formate dall'intersezione tra suddivisioni interne della conchiglia, i setti, e la sua superficie interna. I setti suddividono in camere la parte apicale della conchiglia, la cosiddetta camera d'aria o fragmocono, in cui era presente del liquido o del gas mentre il corpo dell'animale si trovava nella parte della conchiglia rivolta verso l'apertura, la camera d'abitazione. Nel corso dello sviluppo, l'animale si spostava verso l'apertura della conchiglia, il cui bordo si accresceva in modo continuo, e formava una nuova camera del fragmocono secernendo un setto per mezzo di un tessuto, il mantello. I setti delle camere del fragmocono sono attraversati da un piccolo tubo, il sifone, nel cui interno si trovavano i tessuti che controllavano il riempimento o lo svuotamento di liquidi o di gas presenti nelle camere.

Le conchiglie dei nautiloidi e delle ammoniti differiscono per la posizione del sifone, per le diverse morfologie del guscio esterno, che nelle ammoniti sono molto più eterogenee, e soprattutto per quel che riguarda i setti, che nelle ammoniti si ripiegano in superfici anticlastiche diventando sempre più ondulati e frastagliati a mano a mano che si avvicinano alla parete interna del guscio dove si inseriscono producendo linee suturali che in genere aumentano di complessità con il procedere dell'ontogenesi e della filogenesi.

Quali meccanismi e principi costruttivi possono aver generato le regolarità geometriche costantemente presenti nell'architettura dell'incredibile varietà di forme delle conchiglie di ammoniti? L'evoluzione dei cefalopodi con conchiglia esterna, come quella di ogni altro organismo vivente, è il risultato dei processi di mutazione e ricombinazione dell'informazione genetica che generano la variabilità su cui possono agire i processi di selezione naturale. Talvolta compaiono individui con caratteristiche nuove che riescono a sopravvivere e a riprodursi con maggiore successo degli altri membri della popolazione, che perciò possono essere soppiantati dai nuovi arrivati. Questi meccanismi evolutivi portano con il passare del tempo allo sviluppo di organismi e organi sempre più complessi, specializzati ed efficienti: si stabiliscono così delle strette corrispondenze tra strutture e funzioni.

Anche la forma dei viventi si evolve in modo da raggiungere una sempre maggiore efficienza funzionale e il massimo grado di adattamento alle forze cui essi sono sottoposti. Perciò si dovrebbe presumere che, come sosteneva il naturalista inglese D'Arcy W. Thompson nel suo libro uscito nel 1917 *Crescita e forma* (edizione italiana Boringhieri, Torino, 1969) «le forme sono un diagramma delle forze». La simmetria bilaterale, presente anche nei cefalopodi, è un buon esempio di una caratteristica comune a diversi organismi viventi la cui affermazione nel corso dell'evoluzione è stata determinata da adattamenti funzionali: per un organismo di certe dimensioni esiste una forte differenza tra sopra e sotto, determinata dalla gravità, e tra davanti e dietro, dovuta alla locomozione, mentre i lati destro e sinistro del suo ambiente presentano gli stessi tipi di stimoli.

Molti ricercatori non sono d'accordo con la pretesa di spiegare tutta l'evoluzione in termini di correlazioni tra struttura e funzione. Alcuni di loro sostengono che l'adattamento è solo una componente secondaria dei processi evolutivi rispetto ad altri meccanismi meno deterministici. Se quest'ipotesi fosse vera, molte caratteristiche ereditarie sarebbero prive di significato funzionale e si affermerebbero o perché «portate a rimorchio» da altre su cui invece agisce la selezione naturale, o per fenomeni casuali di deriva genetica. Processi del genere hanno sicuramente una notevole importanza per quel che riguarda l'evoluzione molecolare, ma è difficile stabilire quale peso abbiano avuto sull'evoluzione della morfologia degli organismi viventi pluricellulari e dei loro organi.

Le fotografie mostrano due esempi delle linee suturali complesse che caratterizzano le ammoniti giurassiche. In alto si vedono conchiglie fossili del genere *Lytoceras*, vissute circa 145 milioni di anni fa nel Titonico, trovate sulla vetta del Monte Acuto nell'Appennino centrale umbro-marchigiano. In questi esemplari la conchiglia esterna è solo parzialmente conservata; dove questa è assente si notano le complesse linee suturali del modello interno, formatosi per la solidificazione dei sedimenti fangosi che avevano riempito le conchiglie. Nell'esemplare in alto a destra è visibile la linea suturale dell'ultimo setto che separava il fragmocono settato, pieno di gas, dalla camera d'abitazione dove si trovava il corpo dell'animale. La fotografia a destra mostra il modello interno di un'ammonite del genere *Calliphylloceras* vissuta circa 165 milioni di anni fa nel Toarciano e trovata sul Monte Catria nell'Appennino centrale umbro-marchigiano. Questi animali con linea suturale complessa, elevato grado di ricoprimento da parte dell'ultimo giro e camera d'abitazione lunga mezzo giro avevano una buona resistenza alla pressione idrostatica, forma idrodinamica e assetto stabile e quindi dovevano essere buoni nuotatori, adattati a vivere anche a notevoli profondità.

Questo perché spesso è difficile decidere se non esiste un significato funzionale per una certa struttura oppure se, più semplicemente, tale significato ci è sconosciuto a causa della limitatezza delle nostre conoscenze.

Il ruolo dell'adattamento nell'evoluzione morfologica della conchiglia dei nautiloidi e delle ammoniti è stato oggetto di molti studi da parte di numerosi ricercatori e in particolare la questione relativa alla funzione dei setti e delle corrispondenti linee di sutura rappresenta uno degli argomenti più controversi ed enigmatici della paleontologia degli invertebrati. Per esempio il paleontologo Emanuele Padoa scrive nel suo libro *Storia della vita sulla Terra* (Feltrinelli, Milano, 1959) che «senza dubbio nelle sue linee generali l'evoluzione è andata da forme a sutura semplice verso forme a sutura complessa», ma non riesce a decidere se «le complicazioni dei margini settali, le carenature e le nodosità del

guscio rappresentano caratteri su cui ha fatto presa la selezione naturale direttamente, o caratteri correlati, cioè portati a rimorchio da altri, a loro volta favoriti dalla selezione» e conclude affermando che «siamo ancora e probabilmente resteremo sempre all'oscuro del determinismo dei caratteri di sutura dei setti e di forma e ornamentazione della conchiglia delle ammoniti fossili, per potervi appoggiare sopra la teoria razionalmente ineccepibile della selezione naturale».

Secondo Stephen J. Gould della Harvard University esisterebbe una differenza tra fenomeni di microevoluzione, limitati alle popolazioni, nei quali agisce la selezione naturale, e fenomeni di macroevoluzione, nei quali la selezione naturale ha un'importanza secondaria rispetto ad altri processi come quelli che portano a un rapido ed elevato ritmo di speciazione. Gould considera l'evoluzione della linea di sutura delle ammoniti una prova a favore di questa ipotesi. Infatti, secondo Gould, la tendenza della linea di sutura a passare da forme semplici ad altre sempre più complesse non è correlata ad alcun chiaro vantaggio funzionale.

La maggior parte dei biologi e dei paleontologi non è d'accordo con queste affermazioni di Gould. Diversi ricercatori hanno proposto che la struttura dei setti sia utile per l'inserzione dei muscoli alla conchiglia, o per la respirazione, o per il trasporto del liquido camerale, o per aumentare la resistenza della conchiglia alla pressione idrostatica. Quest'ultima ipotesi, formulata per la prima volta da Richard Owen nel 1843, è la più antica e anche la più ragionevole e suffragata dai fatti. Infatti le conchiglie con linee suturali complesse in genere hanno le pareti e i setti più sottili di quelle dello stesso diametro con linee suturali semplici e perciò sono meno pesanti rispetto a queste ultime. Secondo quanto proposto dal paleontologo tedesco E. Pfaff nel 1911 e da Gerd E. G. Westermann della McMaster University di Hamilton nel 1975, l'evoluzione della linea suturale avrebbe contribuito ad aumentare la resistenza della conchiglia alla pressione che l'acqua marina esercitava sull'ultimo setto, dove era trasmessa dal corpo dell'animale, e sulle pareti esterne della conchiglia.

Una notevole quantità di informazioni sulle possibili funzioni della conchiglia dei nautiloidi e delle ammoniti è stata ottenuta da numerose ricerche condotte sui cefalopodi viventi dei generi *Nautilus, Spirula, Sepia* e *Argonauta* che possiedono strutture almeno in parte analoghe a quelle dei loro parenti fossili.

La principale funzione della lunga conchiglia conica diritta dei nautiloidi primitivi che vivevano sul fondo era probabilmente quella di protezione dai predatori. Un lungo cono è però ingombrante e poco manovrabile. Per ovviare a questi inconvenienti è sufficiente avvolgerlo in modo da ottenere una spirale logaritmica. In seguito la parte concamerata della conchiglia, il fragmocono, diventò un importante organo di galleggiamento il cui grado di riempimento con

Vengono mostrati i primi quattro stadi di crescita di una linea di Koch (*a sinistra*) e di un'altra linea simile (*a destra*). A ogni stadio del procedimento al calcolatore vengono ripetute le medesime istruzioni: la parte centrale di ciascun segmento viene sostituita con una punta i cui due lati formano un angolo che è pari a 60 gradi per la linea di Koch e minore per l'altra. Dal momento che ha una punta più stretta, questa seconda linea ha una maggiore dimensione frattale, che corrisponde a un maggior grado di ricoprimento del piano.

gas e liquidi, regolato dai tessuti presenti nel sifone, permetteva di ottenere variazioni della densità dell'animale rispetto a quella dell'acqua marina nell'ambiente circostante. Questo procedimento, ancor oggi utilizzato dal *Nautilus*, consentiva di vivere a qualsiasi profondità e di compiere spostamenti verticali (si veda l'articolo *Il galleggiamento del nautilo* di Peter Ward, Lewis Greenwald e Olive E. Greenwald in «Le Scienze» n. 148, dicembre 1980). Per potersi muovere in acqua mantenendo un assetto stabile è inoltre importante che il punto di applicazione delle forze attrattive gravitazionali, il baricentro dell'animale, si trovi al di sotto del punto di applicazione della spinta idrostatica. La stabilità è tanto maggiore quanto più i due punti sono distanti fra loro. Questo si verifica nelle forme avvolte a spirale quando la camera d'abitazione occupa solo la metà dell'ultimo giro. Altre caratteristiche che aumentano l'idrodinamicità della conchiglia, come l'appiattimento laterale e la presenza di carenature, sono adattamenti per una locomozione più efficiente e, assieme a una buona stabilità, sono caratteristiche delle specie che probabilmente erano buone nuotatrici. Ornamentazioni complesse della conchiglia come coste, spine e tubercoli si trovano frequentemente in forme con assetti poco stabili e potrebbero aver avuto una funzione mimetica o di difesa in specie poco adatte al nuoto che vivevano sui fondali.

Diversi gradi di adattamento al nuoto o all'ambiente in cui le ammoniti vivevano possono aver prodotto l'incredibile varietà di forme e di ornamentazione del guscio esterno di questi animali. In tutti i casi però la conchiglia ha una forma conica, diritta o avvolta a spirale. Il motivo della conservazione di questa particolare geometria dipende dal fatto che la forma conica è una delle poche compatibili con la crescita dell'animale. Per esempio, un guscio a forma di sfera si può ingrandire soltanto attraverso processi di demolizione e ricostruzione che richiedono un inutile dispendio di energia. Altre forme geometriche oltre al cono, come un parallelepipedo o un cilindro, possono accrescersi per aggiunta di nuovo materiale sui bordi, ma solo il cono in questo processo conserva le proporzioni iniziali. Come ha notato D'Arcy W. Thompson «è caratteristico dell'accrescimento delle corna, delle conchiglie e di altre forme organiche in cui è possibile riconoscere una spirale logaritmica che ogni successivo incremento di accrescimento è simile, similmente ingrandito e similmente situato rispetto al precedente ed è di conseguenza gnomone dell'intera struttura preesistente». Il termine gnomone venne creato da Erone di Alessandria che, riprendendo un concetto di Aristotele, definì uno gnomone qualsiasi figura che aggiunta a un'altra conservasse la similitudine fra la figura risultante e quella originaria. Ci imbattiamo a questo punto per la prima volta nel concetto di invarianza di scala: vale a dire la proprietà di un oggetto di possedere la medesima struttura a scale diverse.

Tutte le linee mostrate in questa illustrazione, che hanno dimensione frattale con un valore uguale o vicino a due, possono essere generate al calcolatore attraverso la ripetizione di semplici istruzioni. A ciascuno stadio di crescita da ogni punta si generano due nuove punte di scala sempre più ridotta; esse risultano localizzate nelle regioni centrali della punta preesistente nella linea di Cesaro (*a*), in posizione asimmetrica nella linea di Mandelbrot (*b*) e all'estremità nelle linee con ramificazioni a forma di albero (*c, d*).

Relazioni che dipendono dalla scala hanno profonde implicazioni per la fisiologia degli organismi viventi. Consideriamo per esempio il classico problema, affrontato per la prima volta da Galileo Galilei, dell'incremento della massa di un organismo che cresce al cubo, proporzionalmente al suo volume, mentre la sua superficie cresce solo al quadrato. Un ingrandimento di questo tipo è chiamato ingrandimento lineare perché rispetta le proporzioni fra le lunghezze, ma non conserva le proporzioni tra volumi e superfici. Un sistema che consente di mantenere invariato il rapporto tra volume e superficie, necessario per esempio per gli scambi gassosi, è rappresentato dallo sviluppo di superfici esterne o interne molto irregolari, come il fogliame di un albero o le ramificazioni di un polmone. Ogni volta che un animale aumenta le proprie dimensioni è perciò costretto a risolvere nuove esigenze funzionali.

Anche le ammoniti durante la crescita dovettero affrontare problemi strutturali, soprattutto per quanto riguarda la resistenza del fragmocono pieno di gas alla pressione idrostatica che nell'acqua aumenta di circa un bar ogni dieci metri di profondità.

Un'analisi particolareggiata dei principi meccanici coinvolti nella statica della conchiglia delle ammoniti è stata pubblicata nel 1987 da Roger A. Hewitt, anch'egli della McMaster University di Hamilton, e Westermann.

È possibile valutare la resistenza della conchiglia alla pressione idrostatica utilizzando alcuni principi della scienza delle costruzioni che studia la distribuzione delle forze nelle strutture. Per evitare che una determinata struttura si rompa è necessario che la risultante degli sforzi in ogni punto non superi il limite di deformazione elastica del materiale di cui è costituita. Quindi, maggiore era la quantità di materiale della conchiglia in cui si distribuivano le diverse sollecitazioni e maggiore era la sua resistenza. D'altra parte il fragmocono, essendo un organo di galleggiamento, doveva avere la più bassa densità possibile e perciò doveva ricoprire il massimo volume di gas con il minimo peso di conchiglia. Per realizzare una struttura che, a parità di materiale impiegato, sia la più resistente possibile bisogna fare in modo che gli sforzi siano distribuiti uniformemente in ogni elemento. Questo concetto è espresso da Galileo nei suoi *Dialoghi* per bocca di Sagredo: «Sarebbe una bella cosa se si potesse conoscere la forma da dare ad un solido per renderlo ugualmente resistente in ogni suo punto e per ottenere che un carico, posto nel suo mezzo, non produca una frattura più facilmente che se fosse in qualsivoglia altro punto.» Nel caso delle bolle di sapone o di un campo di forze originato da una carica puntiforme questa forma è quella di una sfera. Invece le ammoniti avevano un guscio a forma di cono avvolto a spirale sottoposto a una pressione di uguale intensità e normale alla sua superficie.

La simmetria radiale della conchiglia ci permette di considerare uno degli assi che uniscono l'apice del cono con un punto situato sul perimetro della sua base come un elemento assiale paragonabile a una trave incastrata agli estremi e sottoposta a un carico uniforme. Questa struttura tende a compensare gli sforzi di trazione trasformandoli in parte in sforzi di compressione. Dato che in un elemento di questo tipo una parte delle tensioni viene scaricata sulle strutture che lo sostengono alle estremità, gli sforzi sono più intensi nelle zone centrali e tendono a fargli assumere il caratteristico ripiegamento ad arco descritto da Eulero. Le sollecitazioni flettenti crescono con il quadrato della dimensione lineare dell'elemento assiale e perciò impongono severe limitazioni alle dimensioni della conchiglia delle ammoniti. Queste limitazioni possono essere superate grazie a una serie di strutture di sostegno perpendicolari alla superficie della conchiglia: i setti. Come un ponte è tanto più robusto quanto maggiore è il numero dei pilastri che lo sostengono, così una conchiglia sarà tanto più robusta quanto maggiore sarà il numero dei setti. Chiaramente oltre un certo limite l'aumento di peso dovuto al materiale impiegato per la costruzione del setto non viene più compensato dal consentito guadagno di materiale sul guscio esterno.

Per riuscire a coprire il massimo volume con la minima quantità di materiale si possono ricoprire i giri precedenti con l'ultimo giro. Un ulteriore miglioramento strutturale della conchiglia si ottiene curvando il setto. Infatti ogni diametro di setto è equivalente a un elemento assiale e tende a incurvarsi. Contribuisce a quest'incurvamento anche la pressione idrostatica esercitata sull'ultimo setto tramite il corpo dell'animale. Inoltre la curvatura del setto permette che parte delle sollecitazioni di compressione normali al guscio esterno venga distribuita tangenzialmente a esso.

In genere la forma della conchiglia è quella di un cono con la base non perfettamente circolare. Ne consegue che la distribuzione degli sforzi sulla superficie del setto è asimmetrica. È possibile calcolare la loro distribuzione semplicemente disegnando uno o più cerchi che meglio approssimino il perimetro del setto: nei punti in cui questi si incontrano, gli sforzi flettenti sono nulli, mentre in tutti gli altri il setto è sottoposto a sollecitazioni che tendono a piegarlo o verso l'apertura della conchiglia o in senso opposto. Si possono visualizzare le sollecitazioni presenti sulla superficie di un setto mediante un polariscopio e una membrana di materiale fotoelastico, della stessa forma del setto, sottoposta a una pressione uniforme sul suo perimetro. Questi sforzi possono essere ridistribuiti ripiegando il setto in una superficie anticlastica complessa. Come ha dimostrato nel 1975 Adolf Seilacher dell'Università di Tübingen mediante simulazioni meccaniche, un setto elastico, libero di deformarsi e scorrere con i bordi lungo la superficie interna di un cono con una sezione ellittica, si dispone in modo da formare una corrispondente linea suturale ad andamento sinusoidale con flessioni dirette verso l'apertura o la parte apicale del cono, chiamate rispettivamente selle e lobi. Linee lobali semplici generate da questo tipo di meccanismi sono presenti nei nautiloidi e nelle ammoniti paleozoiche. La morfologia dei setti di forme con diversa sezione trasversale varia quasi sempre in funzione della distribuzione delle sollecitazioni sul setto. Con l'aumentare delle dimensioni dell'animale si verificava una suddivisione dei lobi e delle selle che portava a un aumento del loro numero e della loro altezza, in modo che lo spazio tra una regione della linea suturale e quelle adiacenti tendeva a rimanere costante.

La struttura vascolare raffigurata a sinistra, ottenuta schiacciando una goccia di materiale viscoso fra due lastrine di vetro e sollevando poi queste lentamente, presenta una notevole analogia morfologica con quella di destra, generata per mezzo di una simulazione al calcolatore. È stato simulato il movimento di 21 particelle disposte inizialmente su una circonferenza a distanze leggermente diverse e in seguito sottoposte a due forze attrattive, una che agiva fra le particelle e un'altra che le attirava verso il centro della circonferenza.

Le simulazioni meccaniche e i metodi analitici di calcolo della scienza delle costruzioni non sono però sufficienti per comprendere l'evoluzione delle complesse linee suturali delle ammoniti mesozoiche. Come vedremo, per questo scopo sono necessari nuovi strumenti matematici e sperimentali.

Ricordo ancora quando da ragazzo durante un'escursione in montagna trovai il mio primo fossile, il modello interno di un'ammonite, e osservai meravigliato la complessa struttura sottilmente ramificata della linea suturale. Cercai invano di scoprire se esisteva qualche semplice regola per generare linee suturali sino a quando non notai per caso, in un vecchio libro di analisi matematica, il disegno di una linea che aveva molte caratteristiche in comune con le linee suturali. Si tratta di una struttura ideata da Helge von Koch nel 1904 che veniva definita una mostruosità matematica, in quanto, pur essendo una linea continua, non può essere derivata in nessun punto. Approfondendo le ricerche su questa linea, nel 1975 venni a conoscenza di un libro appena uscito del ricercatore francese Benoît B. Mandelbrot, che attualmente lavora al Thomas J. Watson Research Center della IBM a Yorktown Heights nello Stato di New York ed è professore di matematica alla Harvard University. In questo libro veniva chiarita la natura della linea di Koch e di altri oggetti matematici simili chiamati «frattali», che hanno in comune la proprietà di essere anche molto complessi e irregolari, ma invarianti rispetto alla scala a cui vengono osservati. Utilizzando i principi della geometria frattale, nel 1978 scrissi un lavoro in cui spiegavo il significato funzionale e il meccanismo di formazione delle linee suturali più complesse. A quel tempo i frattali erano ancora poco conosciuti tanto che non riuscii a convincere quasi nessuno della validità delle mie idee. Fortunatamente negli ultimi anni la situazione è cambiata e vi è stato un notevole interesse verso i fenomeni di crescita frattale soprattutto da parte di matematici, informatici e fisici (si veda l'articolo *L'accrescimento dei frattali* di Leonard M. Sander in «Le Scienze» n. 223, marzo 1987).

La crescita di un frattale può essere facilmente simulata e rappresentata al calcolatore grazie alla ripetizione di semplici istruzioni. Per esempio, la linea di Koch viene generata sostituendo la parte centrale di una linea con una punta costituita da due lati lunghi un terzo della linea originaria che formano tra di loro un angolo di 60 gradi. In questo modo si ottiene una spezzata costituita da quattro segmenti uguali lunghi un terzo della linea originaria. Ripetendo la stessa operazione su ognuno di questi quattro segmenti si ottiene una nuova spezzata più complessa formata da 16 piccoli segmenti, lunghi ciascuno un nono della linea originaria. Si può procedere in questo modo all'infinito ottenendo una linea con caratteristiche intermedie tra un'entità monodimensionale e una bidimensionale. È possibile quantificare questa proprietà dei frattali per mezzo di una grandezza chiamata «dimensione frattale», che nel caso della linea di Koch ci dà un'idea del suo grado di ricoprimento del piano e ha un valore compreso tra uno e due.

Costruendo delle linee simili a quella di Koch in cui le punte sulla parte centrale della linea di partenza sono sempre più strette e formano angoli sempre più piccoli, si possono ottenere frattali con un grado di ricoprimento del piano sempre maggiore. Se la punta è talmente stretta da essere costituita da un segmento, si ottiene una linea, scoperta dal matematico Ernesto Cesàro nel 1905, che

Queste strutture vascolari convergenti sono state ottenute schiacciando gocce di sostanze viscose fra due lastrine di vetro che poi sono state sollevate lentamente. Le strutture vengono prodotte dalle forze attrattive fra le molecole della sostanza viscosa e dalla sua tendenza a collassare lungo il fronte di distacco dai vetri che si sposta verso il centro della goccia. Per ottenere un piacevole effetto estetico e mettere in evidenza le diverse morfologie delle ramificazioni sono state utilizzate sostanze diverse per colore e viscosità.

Lo sviluppo ontogenetico di parti delle linee suturali di *Leptoceras studeri* (*a*) e *Hildoceras lusitanicum* (*b*) del sottordine *Ammonitina*, di *Sowerbyceras calypso* (*c*) del sottordine *Phylloceratina* e di *Lytoceras julleti* (*d*) del sottordine *Lytoceratina* è confrontato rispettivamente con i primi tre stadi di crescita di una linea di Koch (*a*), di una sua trasformazione anamorfica (*b*, *c*) e di una fusione di due linee di Koch una di fronte all'altra. Sono chiare le analogie morfologiche tra le linee suturali delle ammoniti mesozoiche e quelle generate al calcolatore.

ha il valore della dimensione frattale uguale a due. Si tratta di una linea «peaniana», una entità monodimensionale che ripiegandosi ne ricopre un'altra bidimensionale.

Un'altra interessante categoria di frattali simili alla linea di Koch e analoghi alle linee suturali delle ammoniti è stata scoperta da Mandelbrot e viene generata da meccanismi simili a quelli appena descritti, con la differenza che le punte oltre a essere più o meno strette sono collocate su ogni segmento in posizioni asimmetriche. Spostando le punte a una estremità dei segmenti e cambiandone l'inclinazione si ottengono strutture frattali e vascolari a forma di albero.

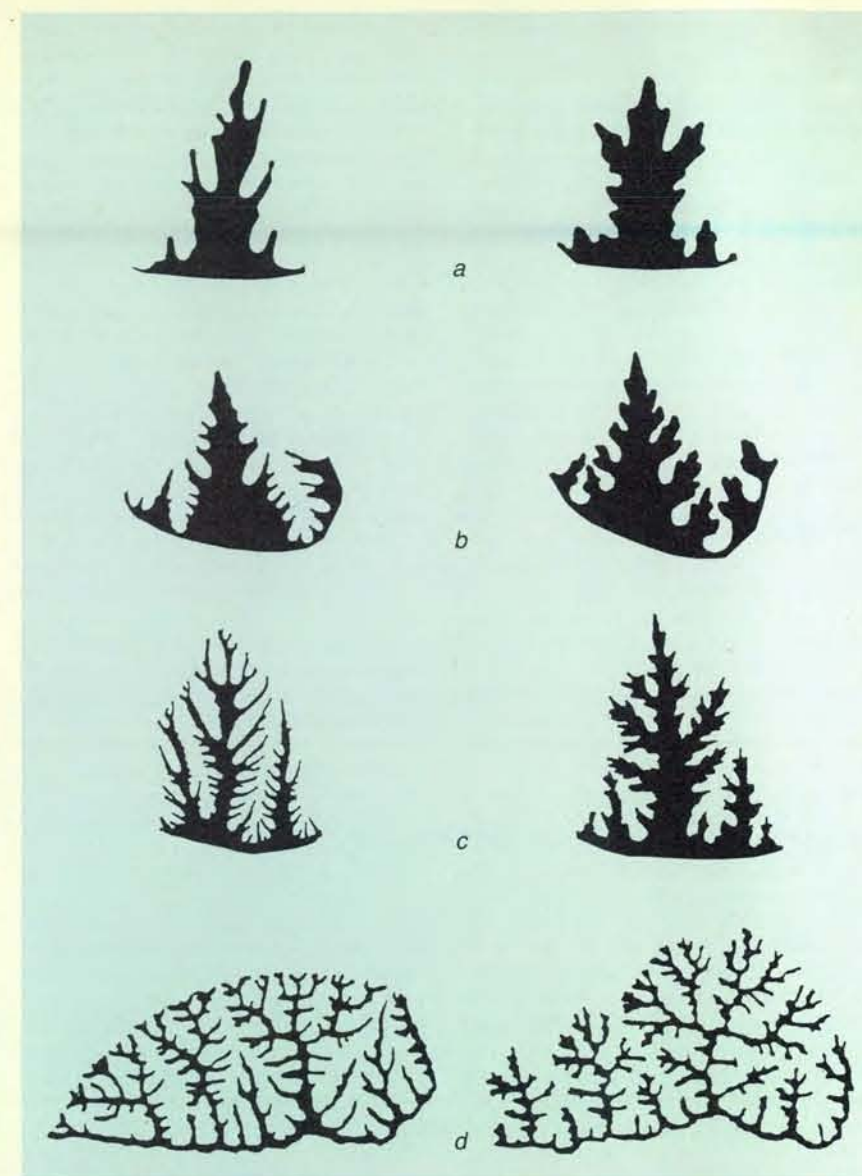
Un altro tipo di programmi per calcolatore che producono frattali analoghi alle linee suturali è costituito dai cosiddetti «automi cellulari». Si tratta di insiemi formati da molti elementi, chiamati cellule, che possono avere solo pochi stati e interagiscono fra di loro. L'evoluzione di un automa cellulare dipende dalla configurazione iniziale delle sue cellule e dalle regole per il calcolo dello stato di ogni cellula a ogni generazione.

Un importante programma di questo tipo è il modello di aggregazione limitata dalla diffusione (DLA) sviluppato nel 1981 da Thomas A. Witten, Jr., della Exxon Research and Engineering Company e Leonard M. Sander dell'Università del Michigan. Questa simulazione al calcolatore parte con un singolo elemento al centro di uno spazio delimitato. Le particelle vengono introdotte nello spazio una alla volta e si muovono casualmente sino a quando si avvicinano abbastanza a un'altra particella da rimanerne legate. Nella struttura che si forma le punte e le zone prive di curvatura hanno una maggiore probabilità delle altre regioni di legare altre particelle e perciò crescono più in fretta. In questo modo si formano frattali irregolari con complesse ramificazioni.

Un altro programma simile al precedente è quello che descrive la formazione di scariche elettriche in una sostanza dielettrica (DBM) sviluppato nel 1983 da L. Niemeyer, Luciano Pietronero e H. J. Wiesmann del centro di ricerca della Brown Boveri & Company di Baden. In questa simulazione la comparsa e la crescita di nuove punte sono regolate da diversi parametri.

Questi automi cellulari generano configurazioni simili a quelle prodotte dalla diffusione di un fluido viscoso sotto pressione in un altro fluido di maggiore viscosità con cui non si mescola. L'apparato per studiare sperimentalmente questi fenomeni è chiamato «cella di Hele-Shaw» ed è costituito da un sottile strato di fluido viscoso, confinato tra due vetri paralleli, al cui centro attraverso un'apertura viene iniettato sotto pressione un fluido meno viscoso. In certi casi questo secondo fluido diffondendosi forma ramificazioni che possono avere diverse morfologie, complessità e dimensioni frattali. Quando come fluido a maggiore viscosità si usa una sostanza nematica a cristalli liquidi che presenta anisotropia a livello molecolare, si verifica una transizione morfologica da strutture con ramificazioni casuali ad altre in cui la crescita di nuove punte avviene in posizioni regolari e prevedibili, dando luogo alla formazione di ramificazioni dendritiche. La simulazione DBM è in grado di riprodurre anche questo tipo di strutture.

Ho scoperto un altro semplice modo per ottenere frattali ramificati, non dalla diffusione di fluidi viscosi, ma dal loro collasso: è sufficiente schiacciare una goccia di una sostanza viscosa fra due lastrine di vetro che vengono poi separate lentamente. Strutture vascolari convergenti vengono prodotte dalle forze attrattive fra le molecole della sostanza viscosa e dalla sua tendenza a collassare lungo il suo fronte di distacco dai vetri che inizialmente ha la forma circolare del perimetro della goccia e si sposta verso il suo centro. Se la velocità con cui si sollevano i due vetri viene aumentata, si formano diverse zone di separazione fra i vetri e la sostanza viscosa che sono distribuite casualmente e quindi in modo abbastanza uniforme all'interno della goccia. Il collasso della sostanza viscosa è localizzato lungo queste numerose zone di distacco che generalmente hanno ognuna un fronte iniziale circolare, provocando una trasformazione della struttura vascolare in un reticolo di poliedri irregolari di forma prevalentemente pseudoesagonale.

Ho sviluppato e studiato molti semplici programmi per calcolatore che producono strutture ramificate, frattali e vascolari che si espandono o si contraggono. Una caratteristica generale di questi programmi è che, quando entità distribuite uniformemente sono diffuse o concentrate da forze repulsive o attrattive, si producono rispettivamente strutture vascolari divergenti o convergenti. Nelle simulazioni che si basano su automi cellulari, la forza repulsiva o attrattiva può essere trasmessa da particolari entità che funzionano da particelle vettori delle forze. Un aspetto molto importante di

Porzioni delle linee suturali (*a sinistra*) di *Silesites seranonis* (*a*), *Perrinites hilli* (*b*), *Kosmaticeras theobaldianum* (*c*) e *Streblites adolphi* (*d*) sono confrontate rispettivamente con ramificazioni prodotte dalla diffusione di sostanze viscose in una cella di Hele-Shaw (*a*, *b*, *c*) o dal loro collasso tra due vetri sollevati lentamente (*d*). In due casi (*b*, *c*) sono presenti ramificazioni dendritiche prodotte dall'uso di sostanze viscose con anisotropie a livello molecolare. La somiglianza tra le linee suturali e le ramificazioni delle sostanze viscose suggerisce l'esistenza di profonde analogie fra i meccanismi che generano queste strutture.

queste simulazioni è che la ripetizione di poche regole semplici produce complesse reti multidimensionali partendo da semplici entità monodimensionali o viceversa.

L'utilizzo della geometria frattale e delle simulazioni al calcolatore conferma la validità dei modelli proposti da Westermann e Pfaff per spiegare l'origine, la funzione e l'evoluzione dei setti e delle linee suturali delle ammoniti. Inoltre questo tipo di analisi rivela nuovi aspetti del processo di costruzione del setto, soprattutto per quel che riguarda l'interazione fra elementi genetici ed elementi ambientali.

Fattori di tipo prevalentemente genetico dovrebbero aver determinato la morfologia del guscio esterno, la viscosità dei fluidi e dei tessuti a contatto con il setto durante la sua costruzione, la sensibilità delle cellule sensoriali agli sforzi e la loro disposizione sulla membrana del mantello che secerneva il setto. Le sollecitazioni sulla conchiglia dovute alla pressione idrostatica erano invece elementi di origine ambientale che contribuivano a determinare la morfologia del setto e della linea lobale.

Si può ipotizzare che il setto venisse costruito sulla base della risposta delle cellule sensoriali a queste sollecitazioni generate dalla differenza di pressione. Le cellule sensoriali probabilmente non erano localizzate solo sulla membrana del mantello, ma anche lungo le pareti dell'ultima camera. Questa ipotesi spiega la relazione esistente fra la morfologia del setto e le dimensioni e la forma del guscio esterno. I gradienti di sollecitazione venivano trasformati in gradienti chimici di sostanze diffusibili secrete dalle cellule sensoriali.

È possibile che in qualche caso la produzione di questi gradienti chimici sia stata determinata geneticamente e perciò in modo indipendente dagli stimoli ambientali. Ma è difficile pensare a un totale determinismo genetico perché in uno stesso animale la linea suturale varia in relazione a cambiamenti nella dimensione e nella forma della conchiglia, come per esempio nel caso della presenza o assenza di coste e tubercoli che provocavano differenze nella distribuzione delle sollecitazioni nella conchiglia e nei tessuti.

In ogni caso le cellule sensoriali in un primo momento si disponevano lungo questi gradienti chimici, corrispondenti ai gradienti delle sollecitazioni, in superfici equipotenziali e in seguito secernevano il setto. Generalmente i valori maggiori degli sforzi erano localizzati nelle regioni centrali con minore curvatura dei lobi, delle selle e dei segmenti che li univano. Perciò queste regioni erano più instabili e sviluppavano nuove punte che portavano alla formazione di lobi e selle più piccoli di secondo ordine. Un'ulteriore risposta delle cellule sensoriali alla nuova distribuzione delle sollecitazioni generava lo sviluppo di nuove punte a una scala ancora più piccola di un ordine successivo e così via. Con la ripetizione di questi processi di ramificazione determinati dalle cellule sensoriali venivano prodotti i setti più complessi e le corrispondenti linee suturali. Queste regole ricorsive sono simili a quelle che generano la linea di Koch e le altre linee e strutture frattali descritte precedentemente: non c'è da meravigliarsi che regole analoghe producano strutture analoghe.

Alcune linee suturali, come per esempio quelle delle ammoniti del sottordine *Phylloceratina*, sono molto convesse e molto allungate rispetto alla linea di Koch. L'allungamento talvolta risulta essere una semplice trasformazione anamorfica della linea di Koch, ma più spesso porta alla presenza di un numero di ramificazioni di secondo ordine sui lati dei lobi e delle selle che aumenta con l'aumentare dell'entità del loro allungamento. Alcune di queste caratteristiche possono essere spiegate considerando la relazione esistente fra i processi di formazione del setto e quelli di diffusione delle sostanze viscose. Le analogie morfologiche delle convessità e delle punte presenti nelle linee suturali e nelle strutture ramificate prodotte dalle sostanze viscose suggeriscono che il fluido presente nella camera in costruzione fosse più viscoso dei fluidi o dei tessuti presenti dall'altra parte del setto, che erano sotto pressione. Inoltre le linee suturali allungate sono analoghe alle ramificazioni anisotrope che si formano quando un fluido poco viscoso e sotto pressione diffonde in un altro più viscoso costituito da una sostanza nematica a cristalli liquidi. Queste linee suturali anisotrope erano probabilmente prodotte quando, durante la formazione del setto, vi era del liquido nematico nella nuova camera. L'esistenza di linee di sutura senza convessità o addirittura con una convessità invertita indica che in questi rari casi la viscosità del fluido presente nella camera in formazione doveva essere eguale o minore di quella delle sostanze presenti dalla parte del setto rivolta verso il corpo dell'animale.

Un'altra caratteristica di alcune linee lobali, come quelle delle ammoniti del sottordine *Lytoceratina*, è quella di essere simmetriche rispetto a un asse longitudinale, a parte le differenze dovute alla convessità. Queste linee suturali probabilmente si formavano quando esistevano situazioni analoghe su entrambi i lati del setto e infatti sono simili a una linea prodotta dalla fusione di due linee di Koch una di fronte all'altra. Questo tipo di linea presenta inoltre il vantaggio di permettere una tassellatura esagonale del piano che è uniforme ed economica. Infatti processi di diffusione che si propagano su una superficie a partire da una serie di punti equidistanti producono un reticolo esagonale, come nei casi delle celle di convezione nella instabilità di Benard e delle strutture prodotte da una sostanza viscosa schiacciata tra due vetri che sono poi sollevati rapidamente.

Un continuo cambiamento nella direzione della curvatura delle linee suturali contribuisce a distribuire più uniformemente le sollecitazioni tangenziali al guscio esterno e infatti le linee suturali, come la linea di Koch, tendono a essere non derivabili. Alcune irregolarità delle linee suturali rispetto a quella di Koch sono dovute al fatto che le estremità delle linee suturali con orientazioni opposte tendono a incontrarsi in modo da permettere la formazione di un arco fra i corrispondenti ripiegamenti sul setto.

In conclusione si può affermare che senza dubbio la costruzione di setti con superfici anticlastiche complesse, secondo i meccanismi proposti, contribuiva ad aumentare la resistenza della conchiglia delle ammoniti alla pressione idrostatica, raccogliendo e distribuendo uniformemente in ogni punto del guscio gli sforzi di trazione e compressione. Durante lo sviluppo ontogenetico e filogenetico delle linee suturali si è verificato un notevole aumento delle loro complessità che ha prodotto un costante incremento della loro dimensione frattale. Un altro aspetto dell'evoluzione delle linee suturali è stata la tendenza a ottenere la riduzione degli spazi fra le suture mantenendo al minimo valore possibile la superficie del setto e quindi la lunghezza della linea suturale. Per questi due principali motivi le linee suturali più complesse ed evolute tendono a essere strutture ramificate, frattali, vascolari, peaniane e non derivabili che ricoprono il piano con una tassellatura esagonale.

Numerose altre strutture con caratteristiche analoghe a quelle delle linee suturali sono comuni in natura: gli alberi, le foglie, i polmoni e più in generale la maggior parte dei sistemi di trasporto e di circolazione di piante e animali di grosse dimensioni, le strutture di aggregazione prodotte dai mixobatteri e dai mixomiceti, i cromatofori di alcuni animali, le cellule nervose, alcuni minerali e cristalli, i fiocchi di neve dendritici, le scariche elettriche, il corso dei fiumi e gli ammassi di materia di grande scala nell'universo.

Un singolo principio di ottimizzazione porta alla formazione di queste strutture ramificate, frattali e vascolari: distribuire o raccogliere qualcosa nel modo più efficiente e uniforme (cioè invariante rispetto alla scala) per mezzo di strutture interconnesse prodotte da semplici istruzioni ricorsive.

BIBLIOGRAFIA

MANDELBROT BENOIT B., *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, 1982.

DAMIANI GIUSEPPE, *Significato funzionale dell'evoluzione dei setti e delle linee di sutura dei nautiloidi e degli ammonoidi* in «Atti del I Convegno Fossili Evoluzione Ambiente», a cura di Giovanni Pallini, Editore Comitato Centenario Raffaele Piccinini, Pergola, 1986.

NITTMANN JOHANN e STANLEY H. EUGENE, *Tip Splitting Without Interfacial Tension and Dendritic Growth Patterns Arising from Molecular Anisotropy* in «Nature», 321, n. 6071, 12 giugno 1986.

HEWITT ROGER A. e WESTERMANN GERD E. G., *Function of Complexly Fluted Septa in Ammonoid Shells. I. Mechanical Principle and Functional Models* in «Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen», 172, n. 1, febbraio 1986.

HEWITT ROGER A. e WESTERMANN GERD E. G., *Function of Complexly Fluted Septa in Ammonoid Shells. II. Septal Evolution and Conclusions* in «Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen», 174, n. 2, marzo 1987.

WEST BRUCE J. e GOLDBERGER ARY L., *Physiology in Fractal Dimension* in «American Scientist», 75, n. 4, luglio-agosto 1987.

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Montagne frattali, piante graftali e altra grafica al calcolatore della Pixar

Posso facilmente immaginare il primo vero lungometraggio generato dal calcolatore. È l'anno 1991. Inciampo camminando nel corridoio tra le poltrone con in mano uno smisurato cartoccio di popcorn sintetico e una bibita contenente alcuni additivi che rendono inutili i normali ingredienti. Le luci si abbassano, il sipario si apre e lo schermo argenteo si anima con una riduzione della trilogia di J. R. R. Tolkien *Il signore degli anelli*. Lo hobbit Frodo passeggia per una stretta valle. In lontananza, i picchi innevati di una montagna si alzano frastagliati verso il cielo. In primo piano, alberi esotici e piante di specie sconosciute brillano alla luce del sole. La scena cambia e compare un mago che scruta in una sfera di cristallo. Al centro della sfera si vede una fortezza con gli spalti merlati avvolti dalle fiamme.

È difficile dire con esattezza quanto potrà essere convincente in questo film Frodo che cammina e parla, ma sono sicuro che le montagne, le piante, la sfera di cristallo e le fiamme riusciranno tutte magnificamente. Il successo sarà dovuto ampiamente al software e all'hardware avveniristici di una società di nome Pixar, già Lucasfilm Computer Graphics Laboratory. Dopo aver visitato questo affascinante centro della grafica al calcolatore di San Rafael in California, posso mettere a parte i lettori dei più riposti segreti di montagne e alberi. Chiunque possieda un calcolatore domestico è ora in grado di generare immagini che assomigliano molto a questi oggetti. La limitazione di spazio che la rubrica impone mi impedisce di trattare estesamente della sfera di cristallo e del fuoco; svelerò, però, i principi di base per generarli.

Nell'ipotetico film descritto, la telecamera potrebbe zoomare sulle cime innevate alle spalle di Frodo. Non si potrebbe vedere un ammasso più spaventoso: ogni vetta è formata da vette più piccole e così via: un regresso all'infinito di piccoli picchi. Perfino un Orco, quella bestia mostruosa dai piedi di cuoio, non si troverebbe a proprio agio su quegli scabri pendii.

In linea di principio è facile generare una catena montuosa di questo aspetto. Per semplificare, ammettiamo che il terreno copra un'area triangolare. Si suddivide allora il triangolo in quattro triangoli più piccoli, trovando il punto di mezzo di ciascun lato e congiungendo i nuovi punti con tre segmenti. Ciascun triangolo viene a sua volta suddiviso nella stessa maniera. Si continua il procedimento fino a raggiungere i limiti di risoluzione o del tempo di calcolo. Il risultato - un reticolo di triangoli piuttosto monotono - può essere ravvivato aggiungendovi un po' di movimento in verticale: ogni volta che si aggiunge alla scena un nuovo punto medio, lo si sposta verso l'alto o verso il basso di una misura casuale. Gli spostamenti casuali, che in generale devono essere ridotti a mano a mano che i triangoli diventano più piccoli, trasformano i triangoli in vette frastagliate che si alternano a valli (*si veda l'illustrazione in alto di pagina 100*).

Perché questa tecnica dovrebbe produrre montagne dall'aspetto naturale? La risposta può risiedere in parte nel fatto che il procedimento produce un frattale: un tipo di oggetto che crescendo rivela un maggior numero di dettagli. A quanto pare, in tutta la natura si possono vedere frattali. Benoit B. Mandelbrot, l'infaticabile studioso di frattali del Thomas J. Watson Research Center della IBM a Yorktown Heights, New York, utilizza linee costiere per illustrare l'idea di fondo. Immaginiamo che ci sia chiesto di misurare la costa francese con un'asta lunga un chilometro. Facendo ruotare l'asta sulle sue estremità, con una faticosa marcia lungo la costa si arriva a calcolare il numero di chilometri. Molte piccole baie e molti piccoli promontori, però, vengono tralasciati e la lunghezza finale misurata in questo modo non è del tutto esatta. Se si ripete l'esercizio con un righello da un metro si ottiene una misura più precisa e più lunga. Anche in questo caso, però, viene trascurato un gran numero di minuscole insenature e lingue di terra. Senza dubbio, un righello da un centimetro sarebbe più preciso.

Come regola generale possiamo dire che la lunghezza della costa misurata aumenta con il ridursi dell'asta di misurazione. La relazione tra la lunghezza misurata e la dimensione dell'asta è un particolare numero detto «dimensione frattale». A differenza di una comune dimensione, una dimensione frattale di solito è espressa sotto forma di frazione, non di numero intero. La linea costiera in questione potrebbe avere, per esempio, dimensione frattale pari a 3/2. Questa forma può essere vista come una via di mezzo tra una forma a una dimensione (una linea retta) e una forma a due dimensioni (un piano). Se una linea costiera fosse relativamente diritta, la sua dimensione frattale sarebbe vicina a 1; se invece fosse molto frastagliata, la sua dimensione frattale si avvicinerebbe a 2 come se cercasse di riempire un piano a due dimensioni.

Il modello frattale della natura implica un regresso infinito di dettagli. Dal punto di vista della grafica al calcolatore, la questione del regresso all'infinito non si pone; è sufficiente che il paesaggio appaia dettagliato a tutti i livelli di ingrandimento. Fino ai limiti di risoluzione dello schermo, le montagne da generare hanno caratteristiche con finezza pari ai triangoli finali usati nella suddivisione descritta in precedenza. Anche se l'algoritmo completo per disegnare montagne è troppo lungo e complesso per poterlo descrivere in questa sede, c'è un semplice programma chiamato MOUNTAIN che disegna il monte Mandelbrot in sezione trasversale. MOUNTAIN illustra l'idea fondamentale di punti di suddivisione a spostamento casuale lungo un asse verticale. L'artista frattale inizia con un unico segmento orizzontale. Si determina il punto di mezzo e lo si sposta su o giù di una misura casuale. Ciascuno dei due segmenti risultanti viene poi suddiviso e perturbato. Il procedimento può essere proseguito in maniera analoga alla tecnica di suddivisione dei triangoli.

MOUNTAIN conserva due matrici, dette *punti* e *linee*, per seguire il profilo montuoso. Ciascuna matrice ha due colonne e un numero di righe adeguato alla risoluzione dello schermo (per esempio 2048). Le due colonne di *punti* contengono coordinate e le due colonne di *linee* contengono indici; ciascuna linea è specificata come coppia di posizioni nella matrice *punti* che designa le coordinate delle estremità della linea. Dato che è interessante osservare come le suddivisioni che si susseguono formino il profilo di una montagna a partire da un poligono apparentemente poco promettente, MOUNTAIN mette ogni generazione sotto il controllo dell'utente. Al termine di un singolo ciclo principale, il programma chiede all'utente se vuole un'altra iterazione. Se la risposta è positiva, l'esecuzione torna all'inizio del programma.

Il ciclo principale trasforma gli insiemi attuali di punti e linee in nuovi insiemi grandi il doppio. Per ottenere questo risultato scorre *linee* una riga alla volta, cerca gli indici dei punti corrispondenti e richiama le loro coordinate da *punti*. Con le coordinate delle estremità di una data linea, il programma calcola le coordinate del punto di mezzo della linea stessa, modificando casualmente, nel corso del procedimento, la coordinata y. L'algoritmo che segue fornisce una base adeguata per un programma. Le variabili j e k indicano le righe di *punti* e *linee* che a un dato momento vengono riempite con gli ultimi risultati della suddivisione. Le variabili pt e ln registrano il numero di punti e il numero di linee che formano la montagna prima che il programma entri nel ciclo principale. All'inizio j è uguale a pt e k è uguale a ln. L'indice i va da 1 a ln.

```
j ← j + 1
k ← k + 1
a ← linee (i,1)
b ← linee (i,2)
x1← punti (a,1)
y1← punti (a,2)
x2← punti (b,1)
y2← punti (b,2)
punti (j,1) ← (x1 + x2)/2
punti (j,2) ← (y1 + y2)/2 +
                    a caso (intervallo)
linee (i,2) ← j
linee (k,1) ← j
linee (k,2) ← b
```

Questa parte di MOUNTAIN si spiega in gran parte da sé. Una volta calcolate le coordinate del j-esimo punto, l'indice j viene memorizzato come secondo punto dell'i-esima linea e primo punto della k-esima linea. Il primo punto della linea i-esima è lo stesso di prima e il secondo punto della linea k-esima è identico al secondo punto originario della linea i-esima, vale a dire quello con indice b.

Al termine del ciclo di calcolo, pt e ln devono essere riportati, rispettivamente, agli ultimi valori di j e di k. Alla variabile *intervallo* inizialmente l'utente dà come valore la massima quantità di casualità verticale che può essere data al punto di suddivisione. Ogni volta che il ciclo viene completato, questa variabile deve essere divisa per 2 in modo che le fluttuazioni casuali siano sempre in scala con la dimensione delle caratteristiche da variare. La funzione *a caso* (*intervallo*) intende esprimere la selezione di un numero casuale compreso tra 0 e il valore di *intervallo* (al momento dato).

Se appaiono efficaci le montagne che stanno alle spalle di Frodo, gli alberi e le piante che lo circondano non sono da meno. Sono allo stesso tempo realistici e fiabeschi. Sembrano reali perché hanno ramificazioni simili a quelle delle vere piante, e fiabeschi perché non appartengono a specie familiari; il progettista grafico ha a disposizione un tale numero di parametri da non poter resistere alla tentazione di creare qualcosa di nuovo.

Le nuove «specie» sono denominate piante graftali, perché sono basate su grafi e hanno un'implicita natura frattale. Per implicita natura frattale intendo il fatto che le regole per generare la topologia di base delle piante potrebbero essere (ma non lo sono) applicate al limite di risoluzione dello schermo. In breve, un ramo non si sviluppa in un regresso all'infinito di rametti. Una volta sviluppato, il grafo che costituisce la base di una pianta può essere trasformato in una miriade di specie convincenti se lo si interpreta in termini di dimensione, colore, spessore, struttura e così via.

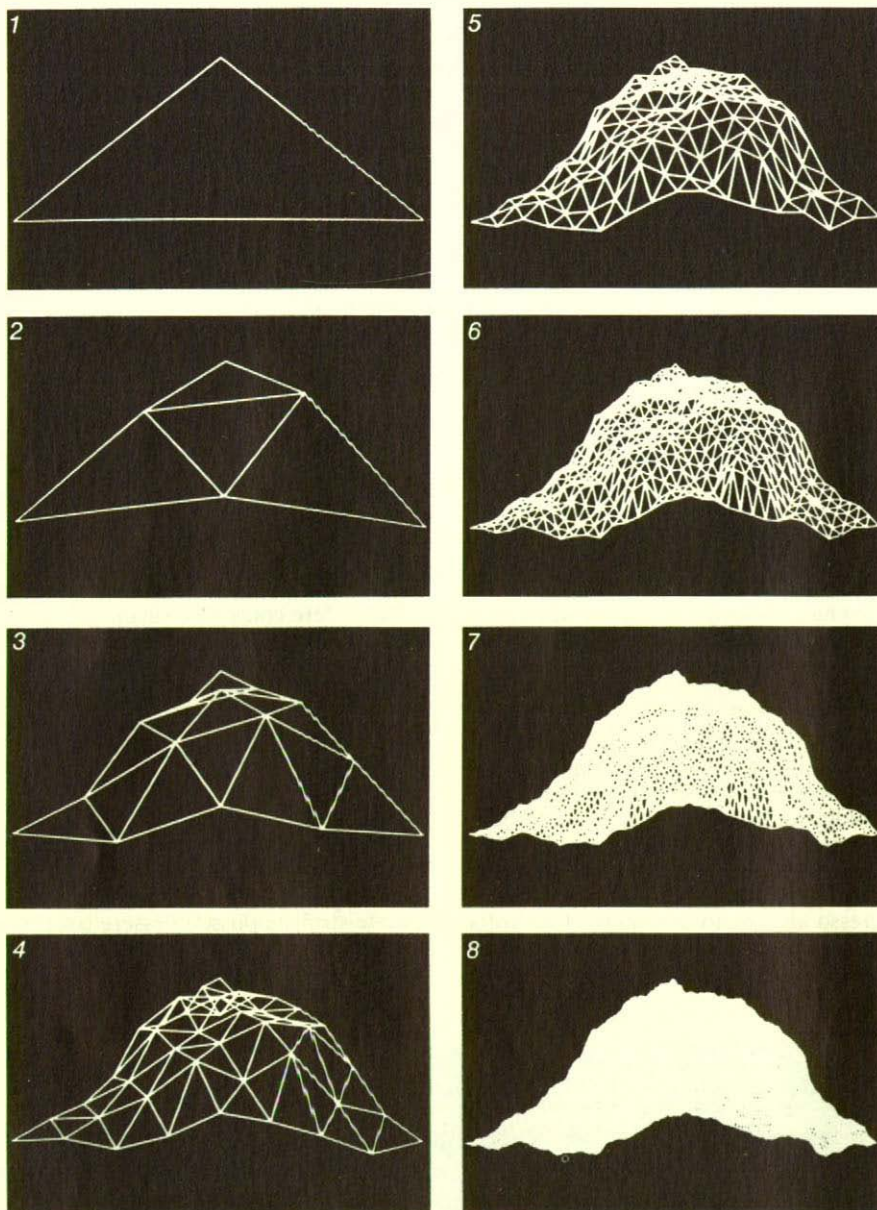
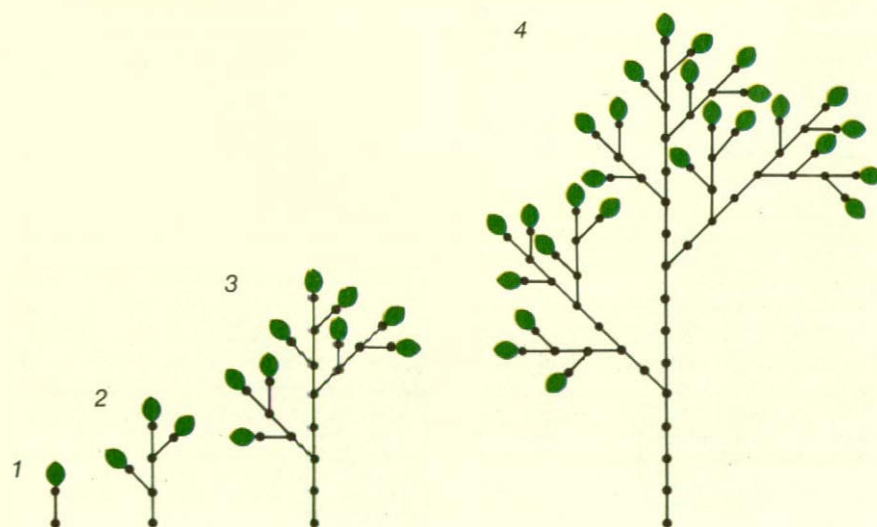
I grafi che sottendono una data pianta sono prodotti da sistemi L, una classe di grammatiche introdotta nel 1968 dal biologo e matematico danese Aristid Lindenmeyer. Un sistema L è in sostanza un insieme di regole per derivare nuove stringhe di simboli da vecchie stringhe. Le regole comprendono sequenze di sostituzione di simboli per singoli simboli. Per esempio, usando i numeri 0 e 1 e i simboli di parentesi quadra aperta e chiusa si può generare una vasta gamma di forme botaniche complesse con le seguenti regole:

```
0 → 1 [0]1[0]0
i → 1 1
[ → [
] → ]
```

Per vedere come funzionano le regole, supponiamo di partire con la stringa formata da un unico 0. A ogni simbolo di sinistra della stringa si sostituisce il suo corrispondente simbolo di destra in modo da ottenere le seguenti stringhe in successione:

```
0
1[0]1[0]0
1 1[1[0]1[0]0]1 1[1[0]1[0]0]1[0]1[0]0
```

Queste stringhe possono essere trasformate in grafi a forma di albero trattando ogni numero (0 o 1) come un segmento

Montagne frattali generate dal calcolatore imitano la natura

La suddivisione di triangoli produce una montagna

Quattro generazioni di una pianta graftale

e ogni parentesi come un punto di ramificazione. I segmenti 0 e i segmenti 1 sono di uguale lunghezza; tipicamente si distinguono perché si lasciano spogli tutti i segmenti 1, mentre si pone una foglia all'estremità esterna di ogni segmento 0.

Per esempio, lo stelo (o tronco) della stringa 1[0]1[0]0 è formato dai tre simboli che non si trovano tra parentesi; un segmento 1 è sormontato da un secondo segmento 1 e poi ancora da un segmento 0. Due rami, ciascuno formato da un singolo segmento 0, germinano da questa formula. Il primo ramo ha l'attaccatura sopra il primo segmento e il secondo sopra il secondo segmento. Prima di studiare l'illustrazione in basso a sinistra, i lettori potrebbero divertirsi a disegnare un po' di generazioni della struttura. Per amore di realismo, si possono aggiungere al modello altre caratteristiche interpretative. Si potrebbe specificare che per qualunque tronco dato (indipendentemente dal fatto che sia o meno il tronco principale) i rami dovrebbero dipartirsi alternativamente verso destra e verso sinistra. Non volendo imporre alcunché di più arduo ai lettori che desiderassero programmare graftali, sono lieto di suggerire, per le piante, semplici ramoscelli. I professionisti della Pixar trasformano la grammatica appena descritta in bellissime piante come quelle dell'illustrazione in alto della pagina a fronte.

Un programma in due parti detto PLANT genera l'n-esima stringa della successione precedentemente illustrata e poi la presenta come disegno al tratto. Nella sua prima fase, PLANT conserva le stringhe che genera in due matrici di simboli dette *stringaA* e *stringaB*. Ciascuna generazione di piante occupa una delle due matrici in modo alternato: la generazione di una matrice è derivata dalla generazione precedente dell'altra. Non è strettamente necessario immagazzinare simboli in queste matrici: se il programma esegue correttamente le sostituzioni, andranno benissimo anche i numeri 0, 1, 2 e 3.

Le regole del sistema L sono inglobate in enunciati condizionali. Per esempio, si può adattare il seguente passo di codice algoritmico per trasformare uno 0 nella i-esima posizione di *stringaA* in nove nuovi simboli di *stringaB*:

```
se stringaA(i) = 0, allora
    stringaB(j) ← 1
    stringaB(j + 1) ← 2
    stringaB(j + 2) ← 0
    stringaB(j + 3) ← 3
    stringaB(j + 4) ← 1
    stringaB(j + 5) ← 2
    stringaB(j + 6) ← 0
    stringaB(j + 7) ← 3
    stringaB(j + 8) ← 0
    j ← j + 9
```

Qui 0 e 1 stanno per se stessi, mentre 2 e 3 stanno, rispettivamente, per [e]. Se l'i-esimo simbolo di *stringaA* è 0, allora il programma inserisce la successione 1, 2, 0, 3, 1, 2, 0, 3, 0 in nove posizioni successive della matrice *stringaB* a partire dall'indice j (la prima posizione della seconda matrice che non è stata ancora riempita). Un unico ciclo nella prima fase di PLANT contiene quattro enunciati condizionali di questo genere, uno per ogni possibile simbolo incontrato. Il ciclo usa l'indice j come riferimento al simbolo della generazione che si sta elaborando. Il ciclo viene eseguito per il numero di generazioni voluto dall'utente. A ogni stadio, PLANT può chiedere all'utente se vuole un'altra (più lunga) stringa di simboli.

La seconda fase di PLANT, quella grafica, trasforma in un disegno la stringa prodotta dalla prima fase. L'operazione è compiuta ricorsivamente. Finché non incontra una parentesi sinistra, o 2, disegna una successione di segmenti in una data direzione. Quando viene presa in esame una parentesi sinistra di una data coppia, il programma disegna il successivo segmento in una nuova direzione, spostata di 45 gradi in senso antiorario rispetto a quella precedente. La fine del procedimento è segnalata dalla comparsa della corrispondente parentesi destra; qui può essere disegnata una foglia (di forma e colore completamente affidati alla fantasia del lettore). La comparsa di una seconda parentesi sinistra provoca la ripetizione del procedimento, solo che ora la nuova direzione è di 45 gradi in senso orario. Il resto è automatico.

PLANT utilizza un fattore di scala che dipende dalla complessità della pianta da disegnare. L'n-esima generazione, per esempio, è alta approssimativamente 2^n segmenti. Se lo schermo ha un'altezza di 200 pixel, i segmenti devono essere più corti di $200/2^n$. Senza dubbio, i lettori ambiziosi cercheranno varianti nella grammatica generativa, negli angoli di ramificazione e nella forma delle foglie. Se si eseguono queste varianti sullo stesso schermo, appariranno paesaggi di piante e alberi (non molto realistici, bisogna ammettere).

La sfera di cristallo dell'ipotetico film tratto da Tolkien sarebbe realizzata con una tecnica chiamata tracciamento di raggi (*ray tracing*); i merli in fiamme sarebbero simulati seguendo il movimento di un grosso sistema di particelle.

Il tracciamento di raggi richiede di specificare sia la geometria tridimensionale di una scena sia la posizione di una sorgente di luce. Quando lascia una sorgente, la luce si imbarca in una complicata storia di riflessioni e rifrazioni. L'occhio di un osservatore che si trovi sulla scena intercetterà alcuni raggi di luce che fluiscono dalla sorgente ma ne mancherà molti altri, in realtà la maggior parte. Per non sprecare tempo e potenza di calcolo, la tecnica del tracciamento di raggi lavora nella direzione opposta. Immaginiamo per un momento che la luce lasci invece l'occhio. Un ampio fascio di raggi si distende a ventaglio nella scena. Se un raggio colpisce una superficie riflettente o rifrangente, saetta via in una nuova direzione determinata dalle leggi dell'ottica. Infine il raggio colpisce una superficie assorbente, assumendo il colore ivi assegnato. Quel colore è registrato nel pixel corrispondente alla direzione del raggio di partenza.

Piante graftali prodotte alla Pixar

Un'immagine di palle da biliardo generata dal calcolatore illustra la tecnica di tracciamento di raggi

Qui sopra si può vedere un'immagine generata con questa tecnica. I raggi tracciati nella scena composta da palle di biliardo hanno una storia relativamente semplice. Nonostante la semplicità, però, possiamo vedere riflessi l'interno di una sala da gioco e un uomo che, in piedi con una stecca in mano, osserva il colpo.

Il grande sistema di particelle che si potrebbe usare per generare i merli in fiamme è lo sviluppo logico dei piccoli ammassi di punti che rendono simbolicamente le esplosioni in miniatura nei videogiochi. Alla Pixar, però, un sistema di particelle è molto più raffinato. All'interno di una certa regione, moltissime particelle vivono, si muovono e conducono la loro esistenza. Sotto il controllo del calcolatore, ciascuna particella è un punto che si sposta secondo una dinamica predeterminata. Nata in un certo istante, può muoversi per un po', magari dando anche vita a nuove particelle; poi può morire.

Sistemi di particelle sono stati usati in modo spettacolare in una scena del film *Star Trek II*. Una bomba Genesi viene sganciata su un pianeta morto pieno di crateri. La bomba crea un anello di stra-

Una scena dalla sequenza Genesi in **Star Trek II**

ne fiamme scintillanti che alla fine inghiottono il pianeta. Quando infine si spengono, vediamo la superficie trasformata in una lussureggiante biosfera. L'effetto è stato prodotto dalla Pixar nella sua precedente incarnazione come Lucasfilm Computer Graphics Laboratory. L'anello di fuoco in espansione era formato da sistemi di particelle in cui alcune particelle davano origine a interi nuovi sistemi. I nuovi sistemi rappresentavano particelle che, scagliate verso l'alto dalla superficie del pianeta, cambiavano colore e persino ricadevano sotto l'influenza della gravità.

Alvy Ray Smith, che dirige il settore ricerca e sviluppo alla Pixar, mi ha accompagnato in un giro dell'azienda durante la mia visita a San Rafael. Smith è già noto ai lettori di «Le Scienze» per il suo lavoro con gli automi cellulari (si vedano le rubriche «Giochi matematici» del maggio 1971 e «(Ri)creazioni al calcolatore» dell'ottobre 1985). Oltre che con Smith, pioniere nell'applicazione del metodo frattale alla vita delle piante col calcolatore, ho avuto un incontro con Loren Carpenter, specialista di montagne frattali, con Robert L. Cook, esperto del tracciamento di raggi, e con William Reeves, creatore dei sistemi di particelle. Nel bel mezzo di una discussione sul software grafico, Smith mi ha sorpreso affermando che l'attività principale dell'azienda non è tanto la produzione di effetti speciali per Hollywood quanto la costruzione di un calcolatore dedicato alla grafica e chiamato, abbastanza naturalmente, Pixar Image Computer.

Al cuore del Pixar Image Computer c'è una memoria da 24 megabyte per 2000 × 2000 pixel. È una risoluzione più che sufficiente per la maggior parte delle applicazioni. Ciascun pixel, per di più, è rappresentato da 48 bit di memoria, sufficienti a conservare copiose informazioni sul colore e la trasparenza. La grande memoria del Pixar è controllata da quattro elaboratori paralleli ad alta velocità, totalmente programmabili, che possono eseguire circa 40 milioni di istruzioni al secondo, una velocità che è di molti ordini di grandezza superiore a quella dei comuni calcolatori. Un'unità video comunica con la memoria a una velocità di 480 milioni di byte al secondo.

I primi Pixar sono stati messi in commercio nel maggio scorso. Sono destinati all'elaborazione di immagini in campo medico, al rilevamento a distanza, al disegno tecnico e all'animazione. Forse saranno usati anche per generare il mio ipotetico film.

La tappezzeria per la mente, argomento dell'articolo del novembre scorso, riguardava immagini al calcolatore quasi, ma non del tutto, ripetitive: un tipo di tappezzeria che ancora non si è visto. In effetti, però, configurazioni simili a queste erano conosciute una generazione fa. Michael Rossman, scrittore e commentatore politico che vive a Berkeley, California, coniò la frase «tappezzeria per la mente» nel 1971. Si riferiva ai delicati disegni che si vedono a occhi chiusi dopo l'assunzione di LSD. Scrive a questo proposito Rossman: «"Sembrava" una tappezzeria, mai a strisce ma con motivi ripetuti: di solito costellazioni di pappagalli, stelle marine, lampadine, fiocchi di neve, innominabili macchie di Rorschach, cangianti con sfrenata creatività... le immagini si trasformavano a catena una nell'altra, come se qualche stampa lineare infinita di Escher con immagini di metamorfosi si fosse animata in un film.»

I lettori ricorderanno che il più semplice tra i programmi per tappezzeria descritti in questa rubrica era stato realizzato da John E. Connett dell'Università del Minnesota. Secondo le sue indicazioni, le coordinate di ciascun pixel vengono elevate al quadrato e sommate. Il risultato è troncato a un intero. Se l'intero è pari, il pixel viene colorato di nero; se è dispari il pixel viene colorato di bianco. La comparsa di una tappezzeria (ripetizioni orizzontali e verticali) può essere spiegata in parte come fenomeno moiré: nella precedente descrizione sono implicite due configurazioni e la loro sovrapposizione crea l'effetto. La prima configurazione è la griglia rettangolare di pixel che costituisce lo schermo di visualizzazione. La seconda configurazione è una serie di anelli concentrici che rappresentano i punti del piano per i quali il procedimento di Connett genera un numero dispari. La tappezzeria nasce come risultato di figure di interferenza ripetitive nelle direzioni orizzontale e verticale. Può capitare che un gran numero di punti consecutivi della griglia cada all'interno degli anelli; il gruppo successivo cadrà quindi all'esterno, e così via. A mano a mano che la distanza dall'origine aumenta, gli anelli diventano progressivamente più piccoli, garantendo così che i centri e i colpi mancati avvengano lungo qualsiasi linea di punti della griglia.

Un sostegno a queste affermazioni può venire dall'esperienza di Paul Braun di Simi Valley, California, che si è fatto vincere dall'impazienza. Lo schermo del suo calcolatore ci metteva così tanto a riempirsi che egli decise di prendere un campione del disegno visualizzando un pixel ogni otto nella direzione orizzontale e in quella verticale. In forma compressa, l'immagine risultante non assomigliava affatto all'originale; Braun, in effetti, aveva cambiato la dimensione della griglia, alterando in questo modo la figura d'interferenza.

Otto Smith di Port Townsend, Washington, ha scoperto la possibilità di modificare notevolmente la configurazione moiré variando semplicemente lo schema dei colori. Come altri lettori, Smith ha scelto formule diverse da quella di Connett. Certe somme o prodotti di semplici funzioni trigonometriche, per esempio, producono immagini agitate e vorticose che ricordano i disegni a inchiostro colorato usati per abbellire le rilegature interne dei libri pregiati. Smith rileva anche che figure moiré e altri effetti di interferenza si presentano con regolarità nella grafica al calcolatore. In questo contesto sono noti sotto il nome di *aliasing*: la tendenza alla comparsa di immagini non volute quando viene digitalizzata una raffigurazione contenente variazioni regolari molto fini.

Mi ha fatto piacere ricevere un «film manuale» in miniatura, che è stato realizzato da Douglas W. Raymond di Orinda, California, ed è formato da minuscole stampe consecutive di tappezzerie di Connet in cui è gradualmente aumentata la dimensione della griglia. Le stampe sono cucite insieme e lo «spettacolo» si ottiene semplicemente facendo scorrere con il pollice le pagine del libretto risultante. Raymond intitola il suo film «Small Bang with Aliases».

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Bellezza e profondità: l'insieme di Mandelbrot e un'orda di suoi cugini detti insiemi di Julia

Fin dalla sua comparsa su queste pagine, nell'ottobre del 1985, l'insieme di Mandelbrot si è rivelato la stella più nuova e più luminosa nel firmamento della matematica popolare. È contemporaneamente bello e profondo. La sua bellezza, in effetti, è solo un velo steso sopra il suo significato: l'osservatore casuale vede solo un groviglio in miniatura di filamenti e svolazzi nei pressi del confine dell'insieme, senza sospettare che questi disegni, in realtà, codificano le varie forme del caos e dell'ordine (*si veda l'illustrazione di pagina 90 in alto*).

Nell'articolo apparso nel 1985 avevo compiuto una semplice incursione nell'insieme di Mandelbrot e rimanevano ancora molte cose da dire. L'insieme è in stretto rapporto con la stabilità e il caos nei sistemi dinamici, rapporto stabilito attraverso alcuni insiemi a esso strettamente correlati detti insiemi di Julia, dal nome del matematico francese Gaston Julia. Gli insiemi di Julia corrispondono a ciascun punto interno (o esterno) dell'insieme di Mandelbrot e hanno anch'essi un'intrinseca bellezza frattale (*si veda l'illustrazione di pagina 90 in basso*). Prima di rivolgere la nostra attenzione a questi insiemi, sarebbe bene rivedere l'insieme che prende nome da Benoit B. Mandelbrot, uno scienziato del Thomas J. Watson Research Center dell'IBM a Yorktown Heights, New York.

L'insieme di Mandelbrot si situa sul piano complesso, che è un piano ordinario con alcuni numeri annessi. Per essere più precisi, ciascun punto del piano complesso è rappresentato da un numero avente la forma $a + bi$. Si possono considerare a e b come le coordinate del punto. La coordinata a è detta parte reale del numero $a + bi$ e la coordinata b è detta parte immaginaria. La i ha anche una funzione di indicatore per aiutare il lettore a distinguere una parte dall'altra. I numeri complessi possono essere sommati addizionando le coordinate separatamente: il risultato è un nuovo numero complesso. Per eseguire la moltiplicazione di due numeri complessi si procede come nell'algebra superiore:

$$\begin{array}{r} 3 + 7i \\ \times\ 2 - 4i \\ \hline 6 + 14i \qquad\quad \\ - 12i - 28i^2 \\ \hline 6 + 2i - 28i^2 \end{array}$$

Perché il risultato sia un numero complesso, il termine $28i^2$ deve essere ridotto usando la proprietà più importante dei numeri immaginari, vale a dire $i^2 = -1$. Così $6 + 2i - 28i^2$ diventa $34 + 2i$. Si può ora presentare la formula chiave, formula che apre la porta all'insieme di Mandelbrot, produce gli insiemi di Julia e, in un certo senso, porta l'ordine nel caos:

$$z \leftarrow z^2 + c$$

Qui z e c sono numeri complessi, ciascuno composto da una parte reale e da una immaginaria. Per elevare al quadrato z e per sommare c si eseguono, rispettivamente, una moltiplicazione complessa e una somma complessa. La formula si anima quando viene iterata, cioè calcolata ripetutamente usando il precedente valore di z per ottenere quello successivo. Ne risulta una successione di numeri complessi che si traduce in una specie di strano balletto sul piano complesso. A ogni iterazione della formula, il più recente numero complesso z si trova a una certa distanza dal suo precedessore e questa distanza è cruciale nel calcolo dell'insieme di Mandelbrot.

Mi piace pensare alla successione dei numeri complessi (punti sul piano complesso) prodotti dalla formula come a vagabondaggi del punto iniziale. Questo numero anela forse all'infinito, anela a danzare per sempre sul piano complesso? Alcuni numeri complessi godono di quel destino. Altri sono confinati per sempre all'interno di una determinata area di forma complicata: si potrebbero chiamare prigionieri, e la loro prigione, l'area di confino, ha pareti frattali.

Nella descrizione precedente ho accennato a una continua ripetizione del processo di iterazione. Ma come si scelgono c e il valore iniziale di z? Una possibilità è dare sempre valore zero a z e scegliere valori diversi per c. Riuscirà a fuggire il prigioniero? Si continua a ripetere l'esperimento, facendo variare sistematicamente c su una porzione del piano complesso. Se il prigioniero fugge, si colora c di bianco; in caso contrario lo si colora di nero. Le pareti della prigione assumono la forma dell'insieme di Mandelbrot. Se, invece di colorare di bianco i punti in fuga, si dà loro un colore che varia con la velocità di fuga, si creano immagini ancora più belle.

Nella regola appena descritta, z iniziava con il valore complesso 0, vale a dire $0 + 0i$. Che cosa succederebbe se si adottasse qualche altro valore di partenza, per esempio $z = 3,5 + 6i$? L'insieme risultante avrebbe una forma diversa? Il risultato è in realtà sempre una versione deformata dell'insieme di Mandelbrot, e allora è preferibile limitarsi all'oggetto canonico.

Seguendo la regola opposta, in cui c è prefissato e z svolge il ruolo di punto iniziale, l'insieme risultante appare assai diverso dall'insieme di Mandelbrot e viene chiamato - o per meglio dire, così vengono chiamati i suoi confini - un insieme di Julia. Avrei preferito dire «l»'insieme di Julia, ma di questi insiemi ve ne sono legioni: per ciascun valore prefissato di c usato nella formula di iterazione, appare un diverso insieme di Julia, pieno di prigionieri.

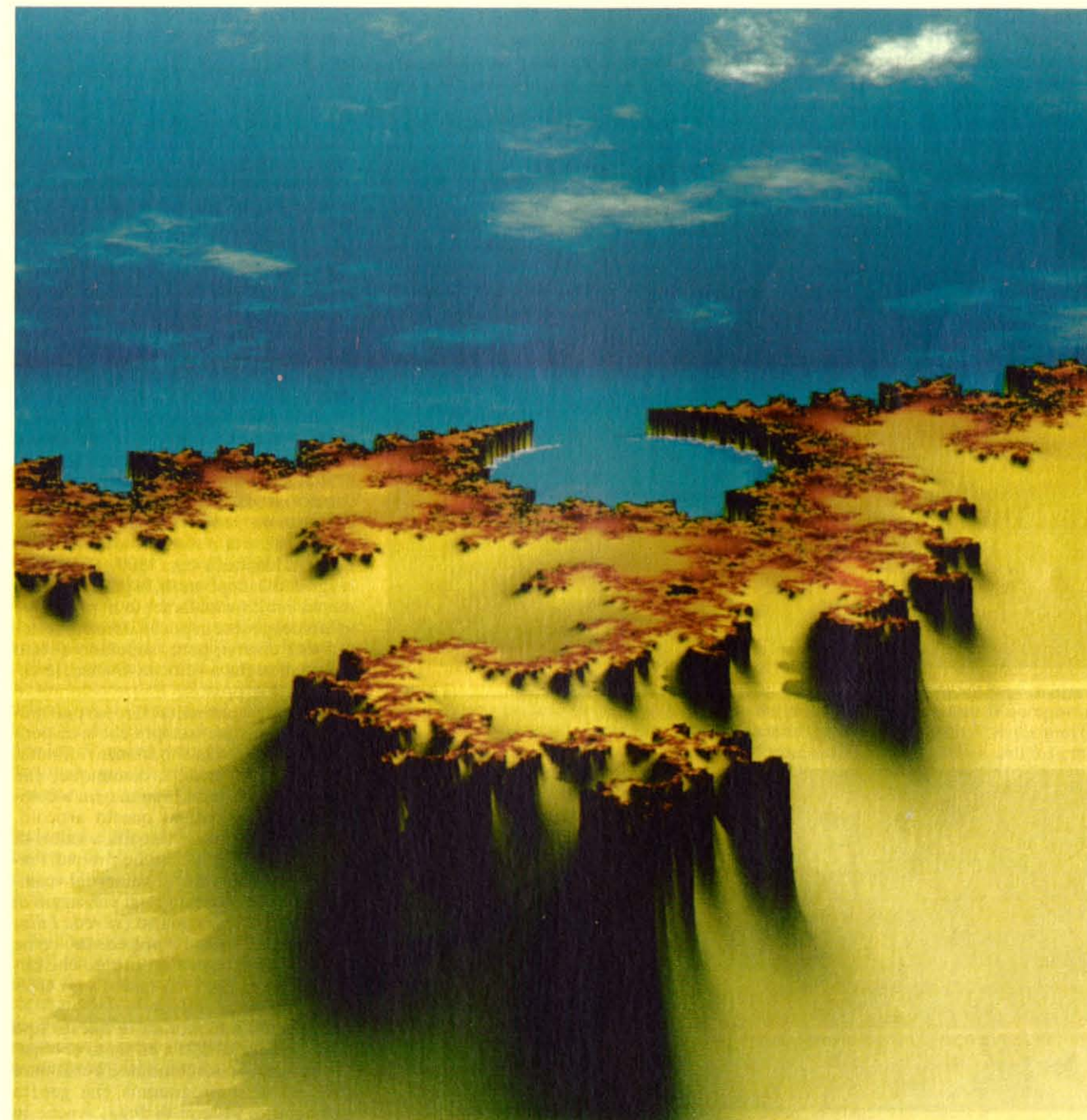
L'ispirazione per questa nuova visita all'insieme di Mandelbrot mi è venuta leggendo *The Beauty of Fractals*, di Heinz-Otto Peitgen e Peter H. Richter, dell'Università di Brema (il volume è stato pubblicato con il titolo *La bellezza dei frattali* dall'editore Bollati Boringhieri). Con le sue splendide immagini in bianco e nero o a colori, oltre che un testo di matematica, è anche un libro da tenere in bella mostra. Le notizie sull'insieme di Mandelbrot, sugli insiemi di Julia a esso collegati e su altri sistemi complessi sono cristallizzate in teoremi a cui di tanto in tanto si aggiungono dei paragrafi con spiegazioni più dettagliate.

Lasciatemi descrivere l'effetto di un teorema. Scrivendo un programma per visualizzare insiemi di Julia, il lettore potrebbe notare che per alcuni valori di c gli insiemi sono ovviamente connessi, ossia costituiti da un unico pezzo, ma per altri valori di c gli insiemi non sono connessi. Che cosa produce la differenza? La risposta è semplice: se il punto c è scelto all'interno dell'insieme di Mandelbrot, il corrispondente insieme di Julia è connesso. Se invece si sceglie c al di fuori dell'insieme di Mandelbrot, l'insieme di Julia non è connesso.

Si potrebbe produrre un incantevole film per illustrare l'applicazione del teorema. Tracciamo una linea retta L da un punto arbitrario interno all'insieme di Mandelbrot a un altro punto esterno e immaginiamo che un punto c si muova in modo lento e costante lungo L all'interno dell'insieme di Mandelbrot e verso il suo confine. L'insieme di Julia associato assume un aspetto sempre più sottile e increspato finché, quando c raggiunge il confine dell'insieme di Mandelbrot, si riduce a un fragile scheletro che non racchiude alcuna area. Quando c supera il confine, il corrispondente insieme di Julia esplode in polvere frattale.

I lettori che abbiano la voglia e la capacità di scrivere un programma possono esplorare l'insieme di Mandelbrot e gli insiemi di Julia inserendo certi algoritmi di base nel linguaggio prescelto. Gli algoritmi hanno in comune il processo iterativo centrale, che dipende fortemente da un particolare teorema: se la dimensione del cosiddetto z iterato raggiunge 2, si perde nell'infinito, senza possibilità di ritorno. Questo fatto, per lo più, distingue i punti in fuga da quelli prigionieri. L'algoritmo lascia a z 100 iterazioni per raggiungere 2. Dato che un numero relativamente piccolo di potenziali fuggiaschi non raggiunge la grandezza 2 in 100 iterazioni, il criterio discriminante non è preciso al 100 per cento. Naturalmente, si potrebbero consentire 1000 iterazioni per ottenere un'immagine un poco più precisa, ma questo richiederebbe una quantità di tempo eccessiva anche su calcolatori veloci.

La grandezza di un numero complesso $a + bi$ è semplicemente la radice quadrata di $a^2 + b^2$ o, in altre parole la sua distanza dallo 0 complesso. Questo è l'algoritmo centrale:

$n \leftarrow 0$
while $n < 100$ e $grand(z) < 2$
  $z \leftarrow z^2 + c$
  $n \leftarrow n + 1$
colora il punto considerato

L'insieme di Mandelbrot sotto forma di

Una variabile indice *n* parte da 0. All'interno di un ciclo «while», che controlla il processo di iterazione, *n* aumenta di 1 a ogni iterazione. Il ciclo «while» fa procedere la formula base fino a che *n* rimane inferiore a 100 e la grandezza di *z* rimane inferiore a 2: se una almeno delle due condizioni viene meno, l'algoritmo esce dal ciclo. Il modo in cui vengono colorati i punti del disegno è affidato al lettore; il colore assegnato dovrà avere ovviamente qualche semplice rapporto di dipendenza da *n*, la lentezza con la quale *z* è fuggito o non è riuscito a fuggire. Il lettore deve anche tenere a mente che il punto dell'immagine è definito da una coppia di coordinate di schermo che saranno diverse dalle coordinate del numero complesso che lì viene tracciato.

Il programma del lettore deve contenere un calcolo separato della grandezza di *z*, rappresentata nel precedente algoritmo da *grand*(*z*). Dato che in genere i linguaggi di programmazione non prevedono un modo diretto per trattare i numeri complessi, *z* deve essere rappresentato in due parti, diciamo *x* (la parte reale) e *y* (la parte immaginaria); *c* deve essere trattato allo stesso modo, magari come *a* e *b*. Il seguente algoritmo sarebbe quindi più simile a un programma funzionante:

```
n ← 0
while n < 100 e x² + y² < 4
    xx ← x² − y² + a
    y ← 2xy + b
    x ← xx
    n ← n + 1
colora il punto considerato
```

I lettori perspicaci avranno notato il trucchetto introdotto in questa versione del processo base: invece di confrontare con 2 la radice quadrata della quantità $x^2 + y^2$, l'algoritmo confronta la quantità stessa con 4. Il risultato è lo stesso e si può evitare di sprecare del tempo ricorrendo continuamente alla funzione radice quadrata. La variabile *xx* conserva temporaneamente il valore appena calcolato di *x* mentre viene calcolato un nuovo valore di *y*. Si salva in tal modo il vecchio valore di *x* per quest'ultimo calcolo, prima di sostituirlo con *xx*.

Si può ora introdurre, in forma un poco più dettagliata, il programma che nel mio articolo di due anni fa ho chiamato MANDELZOOM. L'algoritmo di base è contenuto in un ciclo che fa variare sistematicamente il numero complesso *c* invece delle sue parti *a* e *b*. Se il disegno ha una dimensione di 100 per 100 pixel (punti della griglia), per esempio, ci deve essere un doppio ciclo:

```
gap ← lato/100
a ← verticea
for j ← 1 to 100
    a ← a + gap
    b ← verticeb
    for k ← 1 to 100
        b ← b + gap
        x ← 0
        y ← 0
        [algoritmo di base]
```

Prima di raggiungere queste istruzioni, MANDELZOOM chiede all'utente di introdurre il numero complesso che corrisponde a un vertice del quadrato da esaminare; le sue coordinate sono *verticea* e *verticeb*, i più piccoli fra i valori che *a* e *b* assumeranno nel quadrato. Questo quadrato, specificato dall'utente di MANDELZOOM, dà all'algoritmo il suo nome ed è come una finestra attraverso cui si può sbirciare. Questa finestra può essere resa piccolissima, «zoomando» così sulla parte dell'insieme su cui essa si trova. MANDELZOOM deve anche chiedere all'utente di introdurre un valore per *lato*, la larghezza del disegno sul piano complesso. L'algoritmo calcola poi il divario (*gap*) tra numeri complessi *c* successivi, aumentando opportunamente *a* e *b*.

I valori degli indici *j* e *k* non entrano in alcuno dei calcoli interni al doppio ciclo; è possibile quindi trasformare gli indici in una forma più utile. Per esempio, invece di andare da 1 a 100, *j* e *k* potrebbero variare ciascuno su 100 coordinate di schermo successive. Quando l'algoritmo di base ha deciso che colore assegnare allo *z* iterato, questo colore viene attribuito alle coordinate (*j*, *k*).

Non è possibile abbandonare MANDELZOOM senza ricordare una modifica proposta da Peitgen. Invece che con 2, la grandezza di *z* viene confrontata con 100 o addirittura con 1000. Una volta raggiunto 2, dopo tutto, la grandezza aumenta molto rapidamente e raggiunge valori del genere in poche iterazioni. Iterati di *z* diversi, però, superano questo valore di soglia a velocità diverse; le velocità stesse possono essere colorate, e colorate con continuità! Il rosso può mutarsi in arancione, sempre che si disponga di una tavolozza abbastanza raffinata. È con questa tecnica, comunque, che sono state prodotte le immagini a colori che accompagnano questo articolo. Peitgen paragona le velocità a valori di un campo elettrostatico che circondi l'insieme di Mandelbrot. I valori del «campo» sono rappresentati nel paesaggio di Mandelbrot immaginario (*si veda l'illustrazione della pagina precedente*) come pendii di una catena montuosa che circonda quello che può essere solo chiamato Lago Mandelbrot.

Non sono innamorato di questo tipo di denominazione, ma mi sento obbligato a seguire la consuetudine e a chiamare JULIAZOOM il programma che genera immagini di insiemi di Julia. Anche in questo caso si può zoomare su un insieme per esaminarlo con forti ingrandimenti. JULIAZOOM usa lo stesso algoritmo centrale di MANDELZOOM ma lo inserisce in un contesto un po' diverso.

L'insieme di Mandelbrot (a sinistra), *una sezione del quale* (nel riquadro) *è ingrandita* (a destra)

Insieme di Julia (a sinistra) *corrispondente al centro del disegno in alto a destra; sezione ingrandita* (a destra)

Per prima cosa, JULIAZOOM richiede all'utente *verticex*, *verticey* e *lato*. Richiede inoltre un valore di *c* espresso nei termini delle variabili *a* e *b*. Poi impiega lo stesso doppio ciclo, con alcune notevoli differenze:

```
gap ← lato/100
x ← verticex
for j ← 1 to 100
    x ← x + gap
    y ← verticey
    for k ← 1 to 100
        y ← y + gap
        [algoritmo di base]
```

L'algoritmo di base colora i punti dello schermo a seconda della lentezza con cui le iterazioni raggiungono (o non riescono a raggiungere) la magica soglia di grandezza 2. Alcuni dei grafici di maggior effetto emergono dalle assegnazioni più semplici. Per gli schermi a colori possono dare risultati straordinari anche solo tre colori assegnati secondo lo schema seguente: assegnare il primo colore a punti con valore di lentezza (*n*) da 0 a 10, il secondo colore a punti con lentezza da 11 a 20, il terzo a punti con lentezza da 21 a 30, poi tornare al primo colore per la decina successiva, e così via. Sugli schermi monocromatici si possono ottenere effetti in bianco e nero (o in verde e giallo) usando i due colori e alternandoli di decina in decina.

Sei insiemi di Julia, alcuni connessi (a, b, c *ed* e) *e alcuni no* (d *e* f)

Una volta dotati di una versione funzionante di MANDELZOOM o di JULIAZOOM (o magari di entrambe), i lettori saranno in grado di esplorare autonomamente questi magnifici e significativi insiemi frattali. Si può vagare sul piano complesso in prossimità degli insiemi, oppure zoomare su parti specifiche con il «microscopio per calcolatore» fornito in precedenza. Fino ai limiti di risoluzione consentiti dalla precisione aritmetica della propria macchina, entrambi gli insiemi rivelano dettagli sorprendenti. Per orientare i viaggiatori nel mondo infinitesimale dei frattali, do qui di seguito i domini di coordinate che racchiudono da tutti e quattro i lati entrambi i tipi di insieme:

Insiemi di Julia: *x* e *y* da − 1,8 a + 1,8;
Insieme di Mandelbrot:
x da − 2,25 a + 0,75,
y da − 1,8 a + 1,5

Nell'articolo dello scorso settembre ho affrontato l'argomento del caos in sistemi dinamici quali i pendoli e i circuiti elettronici. L'attività di questo genere di sistemi era riassunta in una semplice formula di iterazione in cui non compaiono numeri complessi bensì numeri reali:

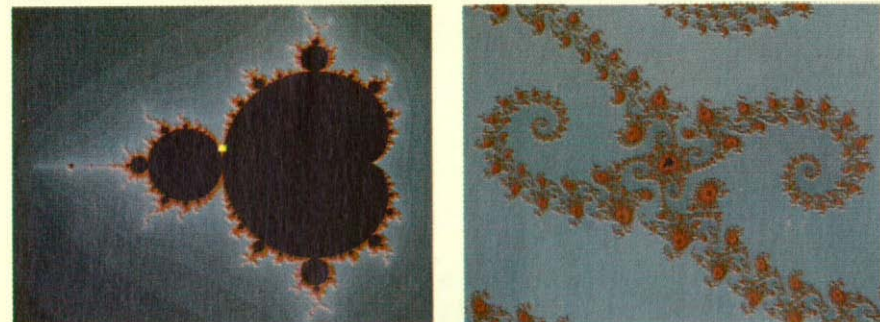
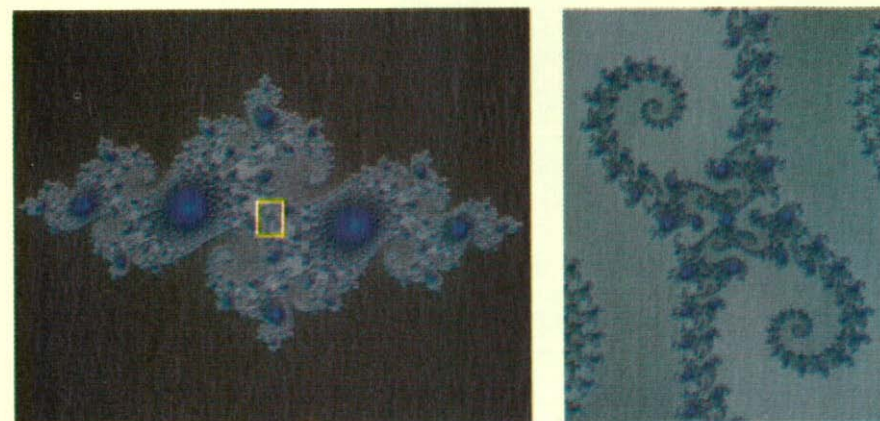
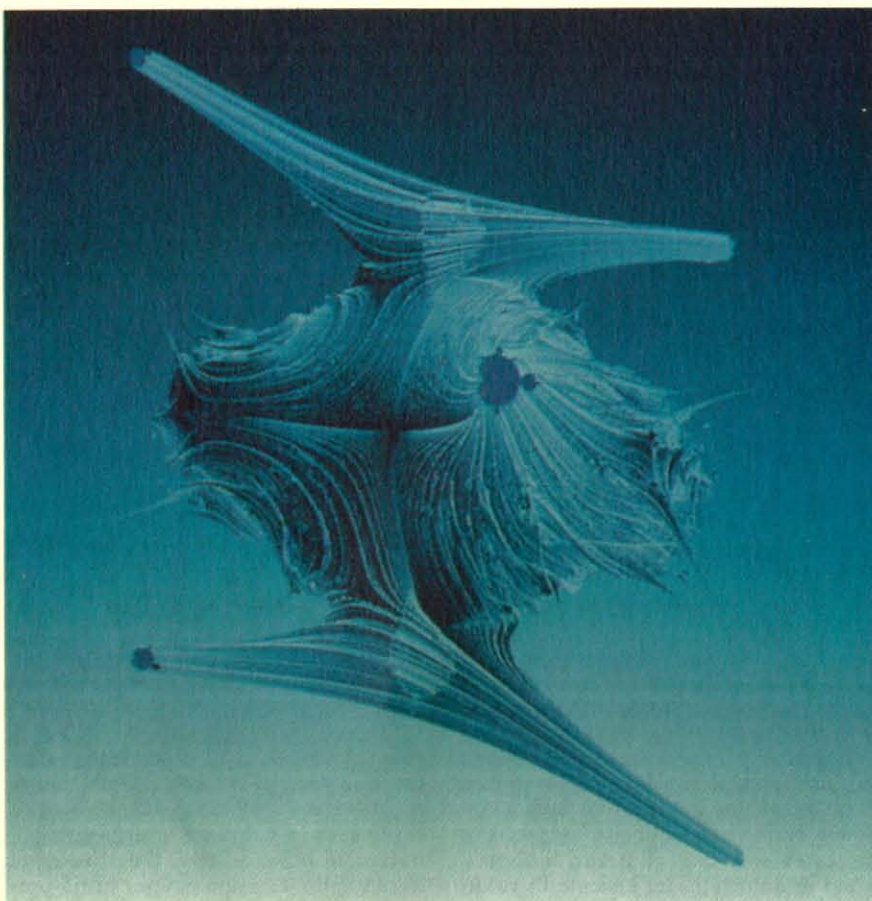
$$x \leftarrow rx(1 - x)$$

La formula è evidentemente quadratica: se si effettua la moltiplicazione, contiene un termine di secondo grado. A seconda di come si sceglie il valore del parametro *r*, quando viene iterata la formula ha un comportamento semplice o strano. Per ciascun valore di *r*, gli iterati si dispongono in un'orbita, un insieme di valori che *x* visita sistematicamente. A un valore critico prossimo a 3,5699, gli iterati hanno oscillazioni ampie e più o meno imprevedibili tra una moltitudine di valori. Questo comportamento corrisponde alla situazione in cui il sistema sottostante, sia esso un pendolo doppio o un circuito elettronico, risulta completamente bloccato nella sua ricerca di stabilità e vaga follemente in modo del tutto imprevedibile: il caos.

Un fenomeno analogo si produce con la formula complessa di iterazione qui descritta, $z \leftarrow z^2 + c$. Per un dato valore di *c*, però, c'è più di un'orbita d'attrazione, a seconda del modo in cui si sceglie il valore iniziale di *z*. Se *z* ha una grandezza iniziale relativamente piccola, graviterà intorno a un punto specifico. Se il valore è grande, crescerà senza limite: l'attrattore è l'infinito. Il punto specifico e l'infinito stesso costituiscono due orbite d'attrazione puntiformi separate per punti sul piano complesso. Il confine tra i loro domini d'attrazione, l'insieme di Julia stesso, è incredibilmente accartocciato e increspato. È anche un'orbita, ma non d'attrazione in senso tecnico. I punti che già si trovano sul confine vi saltano dentro caoticamente. Non è facile, comunque, calcolare direttamente l'insieme di Julia, innanzitutto perché la precisione numerica di un calcolatore può non consentire di specificare i punti che siano esattamente sul confine; nel corso dell'iterazione la precisione diminuisce e l'iterato vaga nella notte.

Ciascun valore possibile di *c* dà luogo, come ho indicato in precedenza, a un nuovo e diverso insieme di Julia. In un certo senso, l'insieme di Mandelbrot riassume, in un colpo solo, tutti i possibili insiemi di Julia: descrive il destino degli iterati dello 0 complesso per tutti i possibili valori di *c*. Per alcuni insiemi di Julia, la regione caotica è semplicemente una figura a forma di albero o addirittura una spruzzata simmetrica di punti neri. I lettori ricorderanno che questo genere di insiemi di Julia corrisponde a valori di *c* che si trovino sopra o al di là del confine dell'insieme di Mandelbrot.

Ho incontrato di recente Peitgen a una conferenza ad Asilomar, in California. Mentre passeggiavamo chiacchierando lungo la spiaggia, mi ha descritto l'insieme di Mandelbrot come una specie di grande libro di cui ciascun insieme di Julia costituisce una semplice pagina. Dalla posizione di *c* nell'insieme di Mandelbrot, si può prevedere il comportamento generico degli iterati in termini di forma e dimensione globali dell'insieme di Julia associato. In tutto ciò c'è qualcosa di più della semplice connessione. Per esempio, se si sceglie *c* nel «collo» compreso tra il corpo principale dell'insieme di Mandelbrot e uno dei suoi germogli, il corrispondente insieme di Julia risulta «strizzato» a sua volta in colli e germogli. L'analogia che vede l'insieme di Mandelbrot come una specie di dizionario per gli insiemi di Julia implica una differenza fondamentale tra i due tipi di insieme: l'insieme di Julia è simile a se stesso mentre l'insieme di Mandelbrot (anche il suo confine) non lo è. Se fosse simile a se stesso, afferma Peitgen, non potrebbe codificare l'infinità non numerabile di insiemi collegati detti di Julia.

In *The Beauty of Fractals* si possono trovare molte altre cose oltre a quelle qui

Sezione trasversale a tre dimensioni dell'insieme di Mandelbrot a quattro dimensioni

riportate. Sono molto grato a Peitgen per aver fornito le immagini che illustrano questo articolo. Non tutte sono state prodotte nei modi che ho descritto, ma nel libro si possono trovare le indicazioni necessarie.

Come osservazione finale sulle attuali ricerche degli studiosi dei sistemi dinamici, citerò un altro oggetto che si nasconde nello spazio a più dimensioni: un oggetto mostruoso che è il risultato di iterazioni cubiche, con z^3 invece di z^2 nella formula di Mandelbrot. L'oggetto ha quattro dimensioni, con riccioli che spuntano in direzioni inconcepibili. Eppure se ne possono calcolare sezioni trasversali tridimensionali, come quella che si vede qui sopra.

Grazie all'articolo di agosto sugli algorompicapo, continuo a ricevere posta da tutto il mondo. Anche questo mese, come avevo promesso, riprendo l'argomento. Prima di tuffarci nel deserto, per rispondere agli ultimi due rompicapo presentati in agosto, devo correggere un mio errore a proposito dei treni: Manuel Blum, uno studioso di calcolatori dell'Università della California di Berkeley, mi ha fatto notare che la quantità di lavoro necessaria perché un treno ne oltrepassi un altro con l'algoritmo fornito è in realtà proporzionale a n^3 (non n^2), dove n è il numero di carrozze del treno. In parole povere, ciascuna delle n carrozze è spostata n volte di n unità di lunghezza. Blum ha trovato un algoritmo che assolve allo stesso compito con una quantità di lavoro proporzionale a $n^2 \times \log n$. Purtroppo mi manca lo spazio per presentarlo.

Il primo problema della Volpe del deserto riguardava un autocarro in grado di trasportare un bidone di benzina da 200 litri alla volta, oltre a 40 litri di benzina nel serbatoio. Il mese scorso ho dimostrato che, partendo da un deposito in cui erano conservati due bidoni, l'autocarro poteva coprire una distanza massima di 1173 e 1/3 chilometri prima di rimanere a secco.

Chester Nowogorski di Naples, Florida, e Norman Rokke di Wintersville, Ohio, hanno proposto degli algoritmi che indicano come un autocarro, con tre bidoni a disposizione, possa coprire complessivamente 1384 chilometri prima di rimanere a corto di carburante. Molti lettori che si sono cimentati nell'impresa sono rimasti al di sotto di questa cifra. Anche le formule generali presentate da molti lettori si sono rivelate insufficienti una volta sostituito 3 a n, il numero totale dei barili impiegati. Non posso quindi garantire la precisione di formule quale quella di Lawrence Leinweber di Cleveland, Ohio, esempio tipico tra le formule inviatemi che portano alle maggiori distanze:

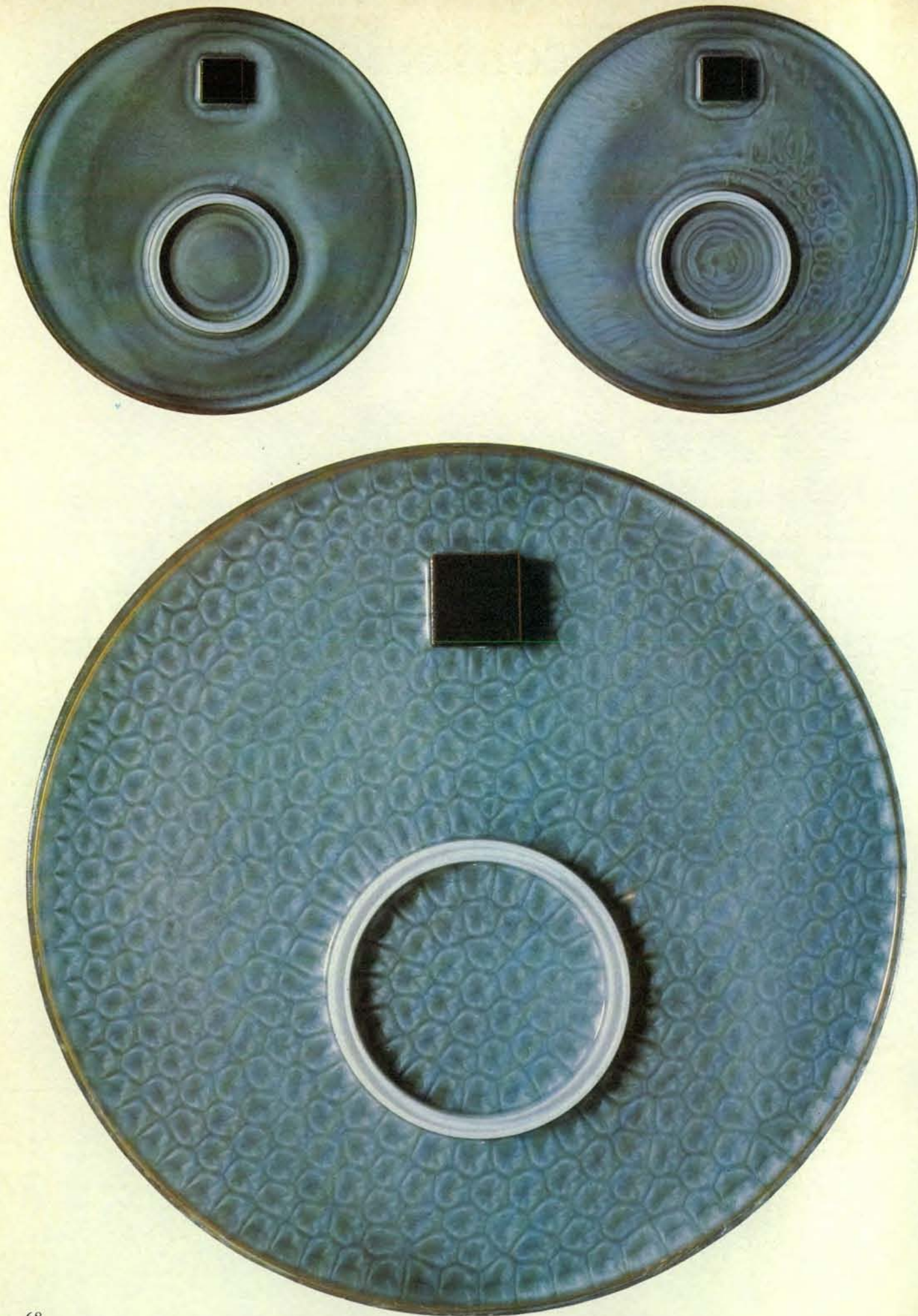
$$5 \sum_{i=1}^{n} \frac{100}{2i-1} - \frac{100}{2n-1} + 100$$

La lettera greca sigma (Σ) è il simbolo di somma: formare n termini con valori consecutivi da 1 a n sostituiti a i in $100/(2i-1)$; sommare gli n termini e poi moltiplicare per 5.

Nel secondo rompicapo della Volpe del deserto, un veicolo di pattuglia può fare rifornimento da n depositi di benzina posti in punti arbitrari lungo un percorso circolare che il veicolo deve compiere. Anche la quantità di benzina di ciascun deposito è arbitraria, ma la quantità totale depositata è esattamente sufficiente a far terminare il suo percorso al veicolo di pattuglia, sempre che quest'ultimo non parta dal deposito sbagliato. Da dove deve partire il veicolo?

Un gran numero di lettori, tra cui Arnold V. Loveridge di Long Beach, California, sono arrivati a un'ingegnosa visualizzazione del problema. Immaginiamo un giro in cui il veicolo di pattuglia parte da un deposito qualsiasi e si avvia in una qualsiasi direzione. Tracciamo un grafico della benzina contenuta nel serbatoio e continuiamo a guidare, anche se il veicolo rimane senza carburante. In questo caso il grafico scende nella zona negativa. A ogni rifornimento a un deposito il grafico si innalza bruscamente e poi inizia un altro lungo e lento declino. Infine il veicolo ritorna al punto d'inizio. Il guidatore deve ora esaminare il grafico e scegliere il deposito dove il veicolo si trovava con meno benzina nel serbatoio (prima del rifornimento). Quello è il deposito da cui il veicolo deve partire.

Numerosi lettori hanno sondato il caos, pronti e desiderosi di inserire in un programma la formula di iterazione fornita nella rubrica di ottobre. In effetti, molti lettori, tra cui Howard Mark di Suffern, New York, hanno accolto il mio suggerimento di seguire il processo iterativo stesso tessendo una tela riflessa intorno a una parabola. In questo caso la convergenza del processo potrebbe essere vista come la tela stessa tessuta in una forma stabile. Ma nel regime caotico riempiva una porzione di schermo tutta di bianco, una confusione di quadrati che non si organizzava in alcuna configurazione distinguibile. Charles A. Plantz di West Brownsville, Pennsylvania, ha usato il suo microcalcolatore come un microscopio, puntandolo sull'arco a forma di staffa del diagramma di biforcazione subito sopra una delle regioni caotiche. Non ha trovato l'alternanza di punti bianchi e neri che mi sarei aspettato, ma pieghe e strati che invadevano la trama del caos.

Gli attrattori strani e la turbolenza

La convezione

Lo spontaneo movimento verso l'alto di un fluido riscaldato può essere compreso soltanto analizzando le complesse relazioni tra temperatura, viscosità, tensione superficiale e ulteriori caratteristiche del fluido

di Manuel G. Velarde e Christiane Normand

La convezione dovrebbe essere familiare a chiunque abbia osservato l'intorbidamento di un brodo riscaldato, sentito il risucchio della cappa di un caminetto od osservato il luccicante tremolio delle correnti d'aria su una strada asfaltata in un giorno di sole. Lo stesso meccanismo convettivo è responsabile delle grandi correnti oceaniche e della circolazione globale dell'atmosfera; esso dà origine a moti su scala ancora maggiore nella fotosfera solare. Certi tipi di nubi si formano quando l'aria calda e umida sale verso l'alto in turbini convettivi ed è l'interruzione del normale trasporto per convezione che lascia periodicamente città come Los Angeles e Madrid sotto una cappa di smog durante un'inversione termica. Altri esempi sono meno familiari o meno facilmente osservati. La convezione ha un'influenza importante sull'asciugatura di pellicole di vernice e sulla dispersione di gas e di particelle nei polmoni. La convezione nel mantello terrestre è in apparenza la forza motrice della lenta migrazione dei continenti.

I più elementari tipi di convezione pare abbiano una semplice spiegazione che si può riassumere con l'espressione «salite di calore». Nei casi più semplici il flusso convettivo ha inizio quando un fluido (un gas o un liquido) viene riscaldato dal sotto. In seguito al riscaldamento lo strato inferiore del fluido si espande e diventa perciò meno denso degli strati sovrastanti. Lo strato inferiore più caldo e più leggero tende quindi a salire, mentre lo strato superiore più freddo tende a scendere in basso. Questo fenomeno era noto fin dal secolo XVIII. Può perciò apparire sorprendente che la formulazione di un dettagliato e quantitativo resoconto della convezione si sia rivelato una duratura sfida per l'ingegno dei teorici. In realtà, perfino di sistemi più semplici sottoposti a un intenso moto convettivo non si può dare una esatta descrizione matematica.

La natura delle difficoltà teoriche può essere esemplificata se si considera ancora il caso di uno strato fluido riscaldato dal sotto. La forza che guida il flusso convettivo in tale fluido è la spinta verso l'alto dello strato riscaldato e il modulo di tale forza è determinato dalla differenza di temperatura tra la superficie superiore e quella inferiore dello strato. La complessità dell'argomento appare chiara quando si scopre che la distribuzione di temperatura è notevolmente alterata dallo stesso flusso convettivo, che trasporta calore dalla superficie inferiore a quella superiore dello strato. In tal modo la forza che provoca il flusso è soggetta a modificazioni da parte dello stesso flusso.

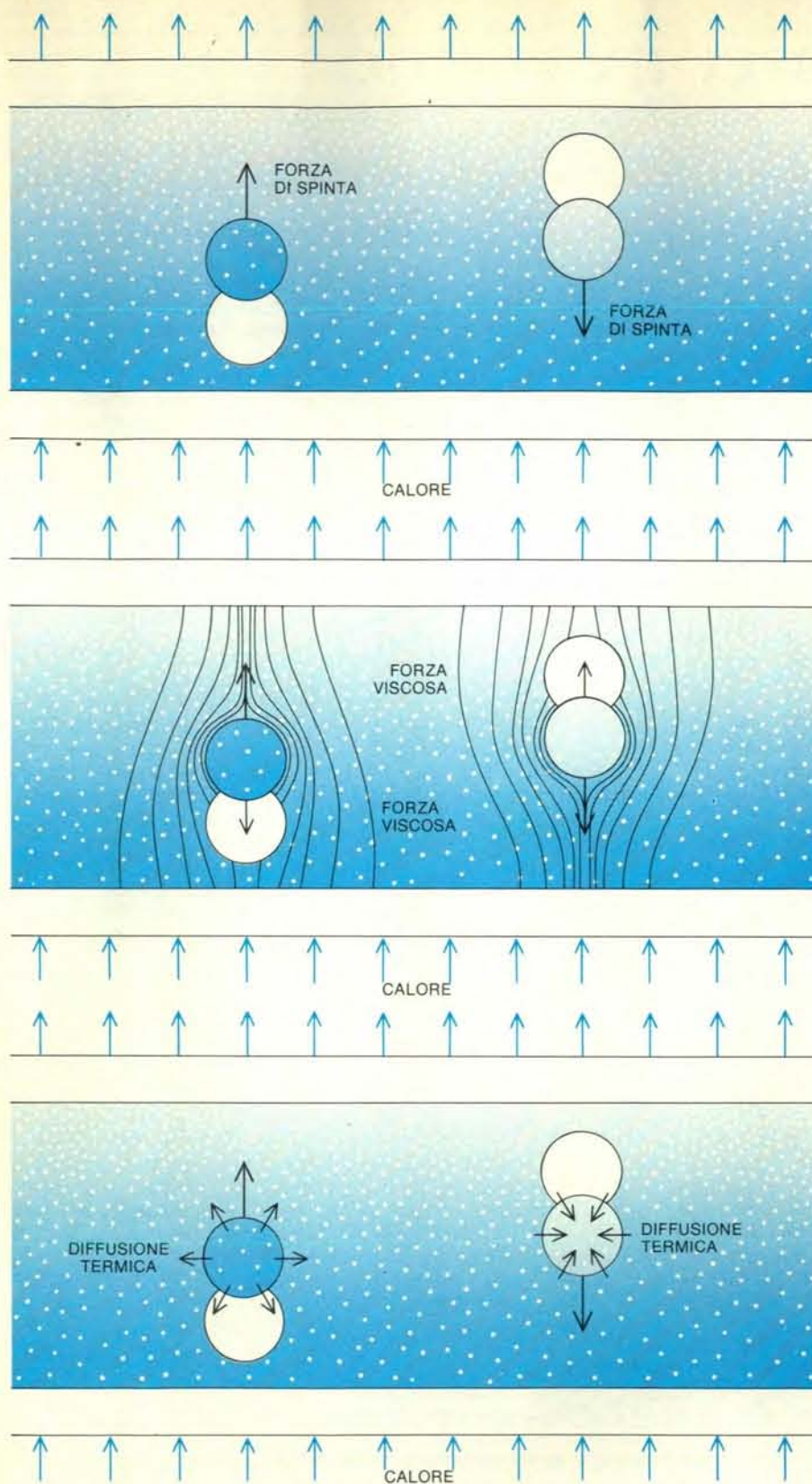
Pur se mancano ancora soluzioni esatte di problemi come questo, negli ultimi vent'anni sono stati compiuti sostanziali progressi verso una teoria generale della convezione. I progressi si sono avuti principalmente dall'adattamento di concetti e di tecniche matematiche sviluppati in altri campi delle scienze fisiche, più che altro nello studio delle transizioni di fase, dei materiali ferromagnetici e dei superconduttori. Questi metodi conducono a una analisi della stabilità delle varie modalità di moto nel fluido e quindi a previsioni della modalità più facilmente osservabile. I risultati sono soltanto approssimati, ma in alcuni casi l'approssimazione è buona, al punto da far sperare che sia possibile spiegare quello che succede in un brodo in ebollizione incipiente.

Il tipo di trasporto convettivo che tratteremo qui è chiamato convezione naturale o libera, che significa che il flusso è dovuto a forze che agiscono in seno al fluido. La forza è molto spesso la gravità, ma vi sono dei casi in cui giocano un ruolo significativo e persino fondamentale altri agenti, quali la tensione superficiale o un campo elettromagnetico. La convezione naturale è definita in contrapposizione alla convezione forzata, nella quale il moto del fluido è indotto da una forza impressa dall'esterno, per esempio da una pompa o da un ventilatore.

Una delle prime descrizioni della convezione naturale risale all'incirca al 1790 ed è dovuta a Benjamin Thompson, il conte Rumford. Egli avanzò l'idea che si poteva spiegare il trasporto di calore in una torta di mele. C'erano state precedenti proposte di un meccanismo convettivo per la circolazione atmosferica e durante il secolo XIX sono stati riportati in letteratura un certo numero di lavori aneddotici. Ricerche sistematiche sono iniziate tuttavia soltanto verso il 1900. Il lavoro sperimentale più significativo di tale periodo si deve al ricercatore francese Henri Bénard. Egli studiò un sistema convettivo più complesso delle sue stesse conoscenze e la sua vera natura è stata riconosciuta solo di recente. Le osservazioni di Bénard e la loro moderna interpretazione verranno riportate più avanti.

Il principale teorico della convezione dell'inizio del secolo XX è stato John William Strutt, alias Lord Rayleigh. Tra i suoi ultimi lavori c'era un articolo sulla convezione, pubblicato nel 1916, che tentava di spiegare i risultati di Bénard. Oggi sappiamo che la teoria di Lord Rayleigh non vale per il sistema esaminato da Bénard, tuttavia il lavoro di Lord Rayleigh è il punto di partenza per quasi tutte le moderne teorie della convezione.

La teoria di Lord Rayleigh può essere spiegata nel contesto di un modello sperimentale che impiega un fluido con pro-

Celle di convezione con una geometria poligonale caratteristica si formano spontaneamente quando viene riscaldato dal sotto un sottile strato di fluido. Nelle fotografie in alto nella pagina a fronte è riprodotta l'evoluzione del modello nel corso di alcune ore; la foto in basso mostra il modello completamente sviluppato. Inizialmente le celle sono lunghi «rulli» che seguono i contorni dello strato; i rulli danno origine a poligoni, che tendono a una forma di esagoni regolari e sono indipendenti dal contorno. In ciascuna cella il fluido sale nel centro e discende alla periferia. La circolazione è indotta principalmente da forze associate alla tensione superficiale e in molti fluidi assume questa forma solo quando la superficie superiore è libera. Le fotografie sono state riprese nel laboratorio di uno degli autori (Velarde) all'Università Autonoma di Madrid. Il fluido è un olio di silicone in cui sono stati sospesi fiocchi di alluminio per rendere visibile il flusso.

Per dare inizio a un flusso convettivo è necessario uno squilibrio di forze. Le forze si possono analizzare in un ipotetico esperimento nel quale un sottile strato di fluido tra due lastre rigide viene riscaldato dal sotto generando un gradiente di temperatura e di densità. Se viene spostata leggermente verso l'alto una porzione di fluido caldo nei pressi della parte inferiore dello strato, essa entra in una regione di maggiore densità media ed è perciò sottoposta a una forza di spinta verso l'alto. Allo stesso modo, se viene spostata verso il basso una porzione di fluido freddo nei pressi della parte superiore dello strato, essa diventa più pesante dell'ambiente circostante e tende a scendere verso il basso. La forza di spinta è ostacolata dalla resistenza viscosa e dalla diffusione del calore, che tende a uguagliare la temperatura di una porzione spostata con quella del suo ambiente circostante. L'importanza relativa di questi effetti è misurata dal numero di Rayleigh. La convezione ha inizio quando la spinta verso l'alto supera gli effetti dissipativi della resistenza viscosa e della diffusione termica, cioè quando il numero di Rayleigh supera un valore critico.

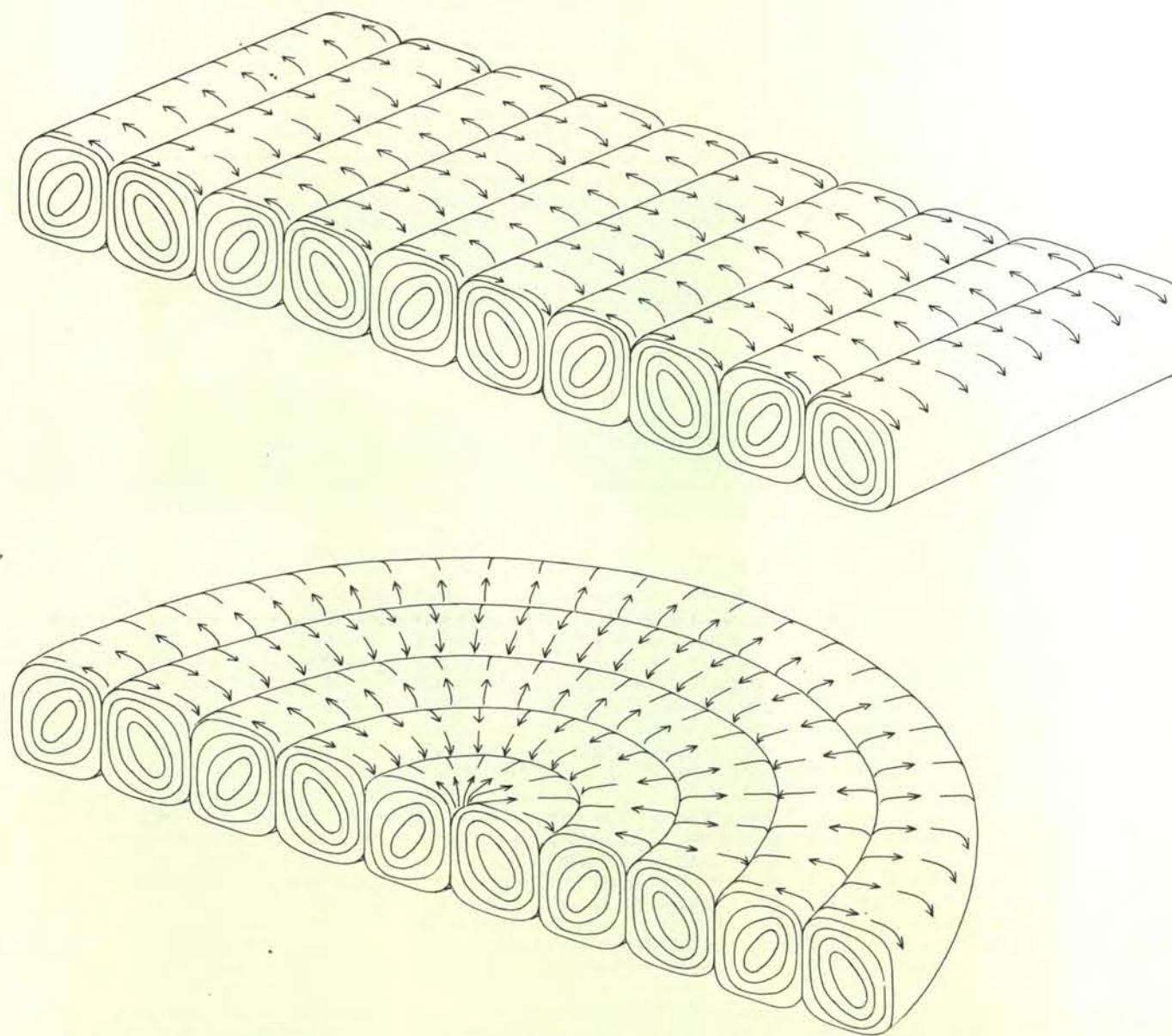
prietà alquanto più semplici di un qualsiasi gas o liquido reale. Un sottile strato di fluido viene confinato tra due lastre piane, rigide, orizzontali e riempie completamente lo spazio tra esse, in modo che non esista alcuna superficie libera. Per strato sottile si intende che l'estensione orizzontale dello strato è molto maggiore della sua profondità (che è uguale alla distanza tra le lastre). L'ipotesi di uno strato sottile serve per rendere minimi gli effetti al contorno delle lastre, che non entrano esplicitamente nella descrizione teorica. Lo strato ideale dovrebbe avere estensione orizzontale infinita; in pratica uno strato di qualche centimetro di larghezza e di pochi millimetri di spessore soddisfa ai requisiti.

L'apparecchiatura deve essere riscaldata dal sotto in modo che il fondo dello strato abbia una temperatura uniforme e costante. Allo stesso modo, viene fornito calore alla parte superiore dello strato in modo che anche là la temperatura sia costante e uniforme, ma più bassa di quella della superficie inferiore. Ne deriva, ovviamente, che anche la differenza di temperatura tra la superficie superiore e quella inferiore è costante e uniforme. Inoltre, si richiede che il gradiente di temperatura - la variazione di temperatura corrispondente a un dato spostamento verticale - sia lineare, cioè che un grafico della temperatura in funzione dell'altezza dia una linea retta.

A questo elenco vanno aggiunte poche altre ipotesi semplificative. Una è che la gravità sia la sola forza che agisce all'interno del fluido. Dato che gli esperimenti pratici sono su piccola scala, il campo gravitazionale dovrà essere praticamente uniforme in tutto il volume. Il fluido deve inoltre essere incomprimibile, ipotesi questa realistica per liquidi poco profondi. Infine, un aspetto molto importante del modello è che la variazione di temperatura alteri direttamente soltanto una proprietà del fluido. Questa proprietà è la densità, che diminuisce con l'aumentare della temperatura. In altri termini, il fluido riscaldato si espande, proprietà questa caratteristica dei gas e dei liquidi reali.

Esiste un metodo teorico ben definito per indagare sugli effetti del gradiente termico in questo esperimento ipotetico. Per prima cosa dobbiamo supporre che una porzione di fluido sia stata spostata verso l'alto o verso il basso rispetto alla sua posizione originale; quindi si devono analizzare le forze agenti sulla porzione spostata. Sono tali forze che determineranno tutti i successivi movimenti del fluido. La porzione può avere dimensioni e forma qualsiasi, ma lo spostamento deve essere piccolo. (La teoria di Rayleigh vale essenzialmente per spostamenti infinitesimali o di entità al limite infinitesime.) Lo spostamento iniziale non deve essere la conseguenza di una forza impressa; dal momento che le molecole di un fluido sono costantemente in movimento le loro posizioni fluttuano casualmente e ci si può aspettare che, se si attende abbastanza a lungo, qualsiasi spostamento piccolo avvenga spontaneamente. Si consideri una piccola porzione di fluido nei pressi della superficie inferiore dello strato. Per l'elevata temperatura sul fondo, la porzione ha una densità minore della densità media dell'intero strato. Però, mentre la porzione resta sul luogo viene circondata da fluido della stessa densità e quindi la sua spinta verso l'alto è nulla. Tutte le forze che agiscono su di essa sono equilibrate ed essa non sale verso l'alto e non affonda.

Supponiamo ora che per effetto di qualche perturbazione casuale alla porzione di fluido venga impresso un leggero moto verso l'alto. Quale effetto ha lo spostamento sull'equilibrio delle forze? La porzione è ora circondata da un fluido più freddo e più denso. Il risultato è che su essa agisce una spinta verso l'alto positiva che la fa salire. La forza complessiva verso l'alto è proporzionale alla differenza di densità e al volume della porzione. Così uno spostamento iniziale verso l'alto del fluido caldo viene amplificato dal gradiente di densità e questo effetto dà origine a forze che provocano un ulteriore movimento verso l'alto. Un'analisi del genere si potrebbe fare per un lieve spostamento verso il basso di una porzione di fluido freddo e denso nei pressi della parte alta dello strato. Muovendosi verso il basso la porzione entrerebbe in un ambiente di densità media più bassa, diventando più pesante dell'ambiente circostante. Essa tende perciò ad affondare, amplificando la perturbazione iniziale. La convezione naturale è il risultato di questi flussi combinati verso l'alto e verso il basso e tende a rovesciare l'intero strato di fluido. La conseguenza di questa analisi è che si dovrebbe osservare la convezione nel fluido tutte le volte che esiste un gradiente termico, per quanto piccolo esso sia. Anche in un gradiente infinitesimo, qualsiasi moto casuale verso l'alto del fluido caldo o verso il basso del fluido freddo dovrebbe essere sufficiente per dare origine a un flusso. In realtà non si osserva tale estrema sensibilità, anzi il gradiente termico deve raggiungere un certo valore di soglia perché inizi il flusso convettivo. La spiegazione del perché le cose stanno così è stato il più significativo contributo di Lord Rayleigh.

Lord Rayleigh mise in evidenza che una teoria della convezione deve tener conto di almeno altri due effetti sul moto di una particella fluida. Uno di essi è la resistenza viscosa, l'equivalente dell'attrito nei fluidi. La resistenza viscosa è sempre opposta al moto del fluido e il suo modulo è determinato in parte da una proprietà intrinseca del fluido, la viscosità di taglio, che misura la resistenza al moto relativo di due qualsiasi regioni adiacenti.

Le celle a forma di rullo costituiscono una configurazione stabile nella convezione indotta da forze di spinta anziché dalla tensione superficiale. L'unità fondamentale del modello comprende due rulli ruotanti in verso opposto; la larghezza di tale unità è doppia della profondità dello strato fluido. La forma del modello dipende dai contorni dello strato. In un recipiente rettangolare i rulli sono paralleli ai lati più corti, mentre in uno cilindrico formano anelli concentrici. Si osserva un modello di rulli stabile soltanto quando il fluido non presenta superficie libera.

Se la velocità non è troppo grande, il modulo della resistenza viscosa è proporzionale alla viscosità moltiplicata per il raggio della porzione e la velocità. Ovviamente, se la resistenza viscosa è uguale alla spinta verso l'alto, non c'è moto.

Il secondo effetto dissipativo costituisce una conseguenza del fatto che la convezione non è il solo meccanismo di trasporto di calore in un fluido. Anche l'irraggiamento e la conduzione, o diffusione del calore, possono trasportare calore. Alle temperature relativamente basse della maggior parte degli esperimenti di convezione, l'irraggiamento fornisce un contributo talmente piccolo che può essere ignorato. Invece la diffusione del calore non è sempre trascurabile; essa tende ad annullare il gradiente termico che origina il flusso convettivo.

L'effetto della diffusione del calore può essere spiegato considerando ancora una porzione di fluido caldo spostata verso l'alto dalla sua posizione di equilibrio in un ambiente più freddo. Secondo la definizione fondamentale di calore, le molecole della porzione calda devono avere una velocità media maggiore di quelle del fluido circostante più freddo. Le molecole possono attraversare liberamente il contorno che definisce la porzione e l'effetto di molti scambi del genere in entrambe le direzioni sarà di uguagliare le velocità medie delle due popolazioni. In altre parole, il calore fuoriuscirà dalla porzione calda spostata in modo da farla raffreddare, mentre il fluido circostante si riscalda, fin quando alla fine essi raggiungono l'equilibrio alla stessa temperatura. Per una porzione di fluido freddo spostata verso il basso, il flusso di calore avviene in verso opposto, dall'ambiente caldo alla porzione fredda. In entrambi i casi la differenza di temperatura locale si riduce e allo stesso modo si comporta la spinta che da essa ha origine.

Il tempo necessario perché una porzione fluida raggiunga l'equilibrio termico con il suo ambiente circostante dipende in parte da una proprietà del fluido, la diffusività termica. La scala dei tempi di questo processo è inversamente proporzionale alla costante di diffusività ed è direttamente proporzionale alla superficie della porzione. Se questo tempo di diffusione termica è confrontabile con il tempo necessario perché la porzione si sposti di una certa distanza caratteristica, quale può essere il suo diametro, la spinta verso l'alto viene annullata. In altre parole, se il fluido si muove non più velocemente di quanto perde calore per diffusione, non può mantenersi un flusso convettivo. La fornitura di calore al sistema attraverso la lastra inferiore viene in tal caso effettuata attraverso lo strato fluido mediante un meccanismo puramente conduttivo o diffusivo, senza alcun moto di assieme.

L'analisi di Lord Rayleigh mostra che la semplice esistenza di un gradiente termico non è sufficiente a garantire l'inizio di un flusso convettivo. È necessario che la spinta verso l'alto risultante da tale gradiente vinca gli effetti dissipativi della resistenza viscosa e della diffusione del calore. L'energia potenziale gravitazionale liberata dalla discesa del fluido più denso e dalla risalita di quello più leggero deve essere maggiore dell'energia dissipata per viscosità e diffusione. La relazione tra tali effetti può essere espressa come un rapporto adimensionale: la forza di spinta divisa per il prodotto tra la resistenza viscosa e la velocità di diffusione del calore. Il rapporto è adimensionale nel senso che tutte le unità di misura associate alle tre grandezze si eliminano esattamente, lasciando un numero puro il cui valore è lo stesso qualunque sia il sistema di unità adottato. Tale rapporto è oggi chiamato numero di Rayleigh. La convezione ha inizio quando il numero di Rayleigh supera un certo valore critico.

Il significato del numero di Rayleigh può essere riaffermato con maggiore esattezza attraverso un esame della stabilità di varie possibili modalità di moto nel fluido. È conveniente definire la stabilità in funzione di una curva di potenziale o di una superficie di potenziale, che esprime l'energia di un sistema in funzione di qualche variabile. Il sistema si trova solitamente in un qualsiasi stato che abbia minima energia, corrispondente al punto più basso sulla superficie di potenziale.

È abbastanza facile immaginare un modello reale di una superficie di potenziale, cioè una boccia emisferica contenente una biglia. In condizioni di equilibrio la biglia è in quiete sul fondo della boccia, dove la sua energia potenziale gravitazionale è minima. Se qualche perturbazione casuale sposta adesso leggermente la biglia, essa rotolerà indietro verso la sua posizione di equilibrio; può darsi che essa superi il punto più basso della superficie e oscilli attorno a esso, ma gli effetti dissipativi dell'attrito finiranno con lo smorzare le oscillazioni e dopo un certo tempo la biglia si ritroverà in quiete nel punto di minima energia. Dato che una biglia sul fondo della boccia ritorna alla sua posizione originale dopo una perturbazione essa è in equilibrio stabile.

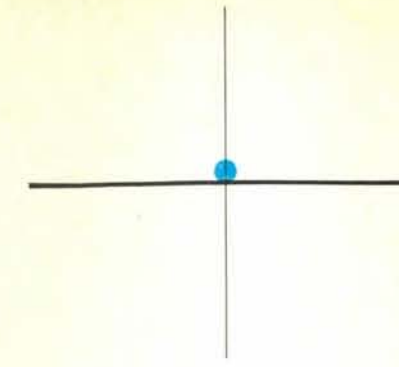
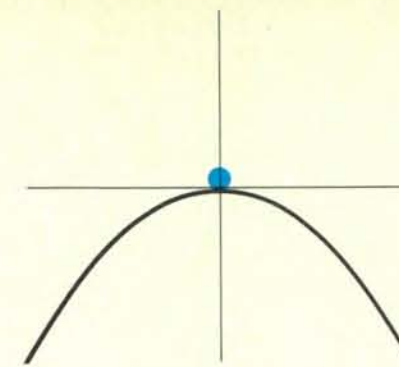
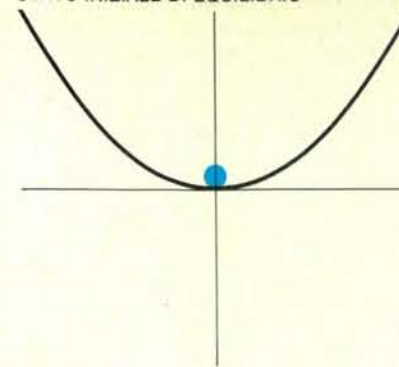
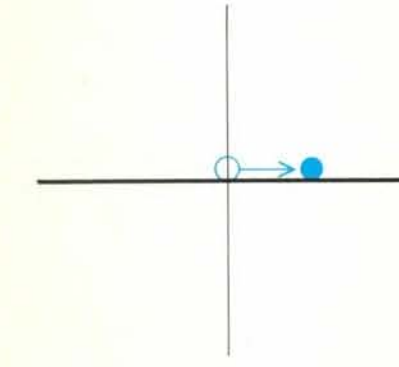
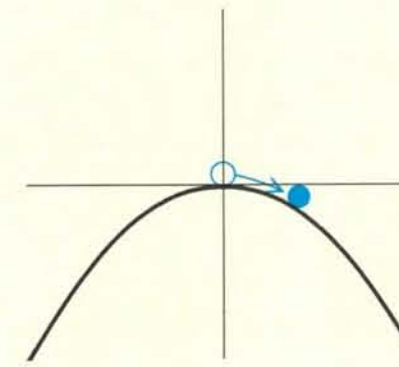
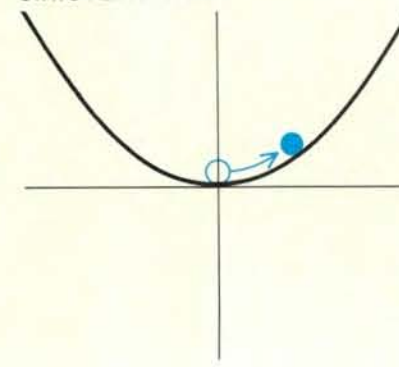
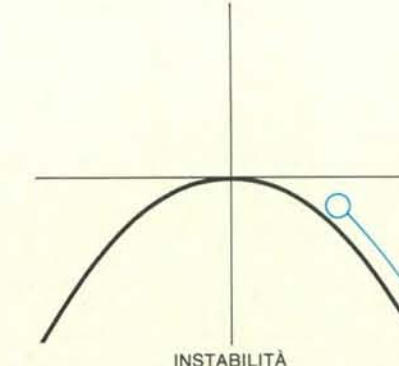
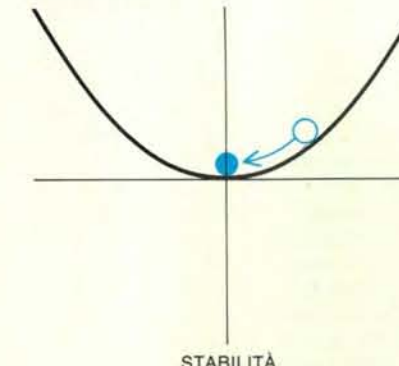
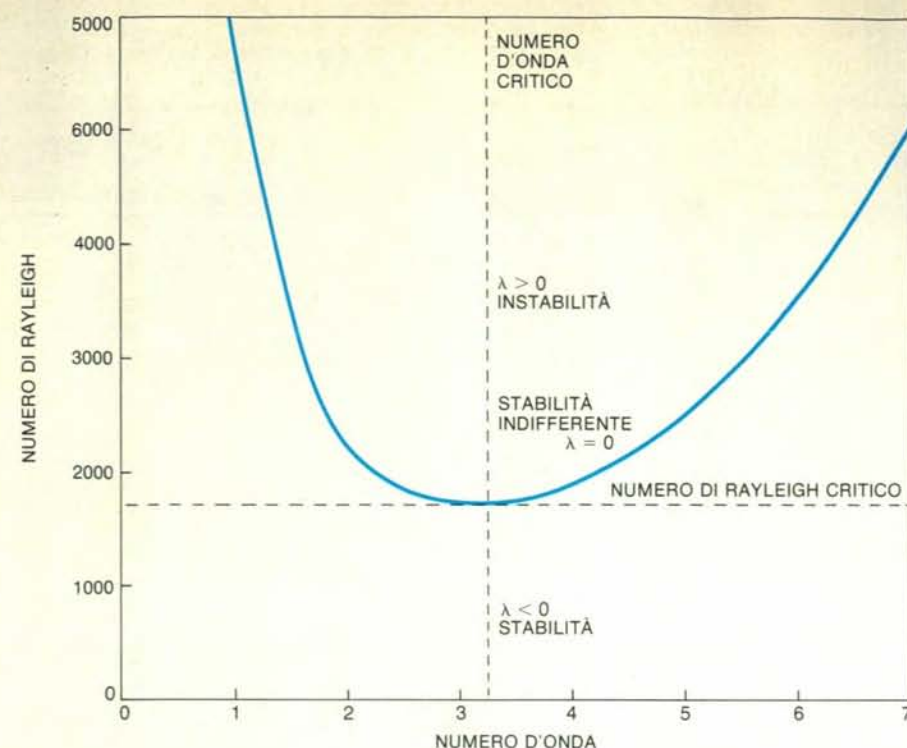
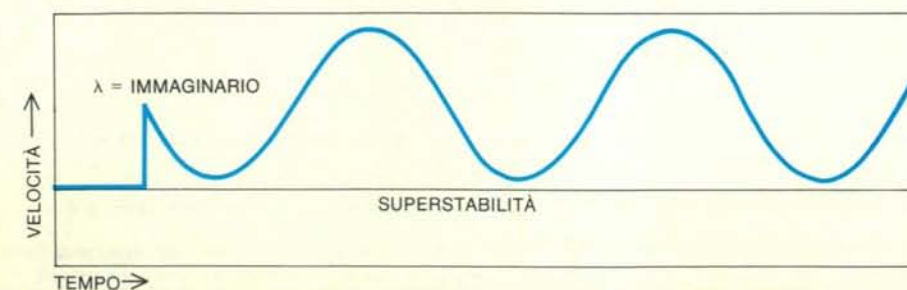
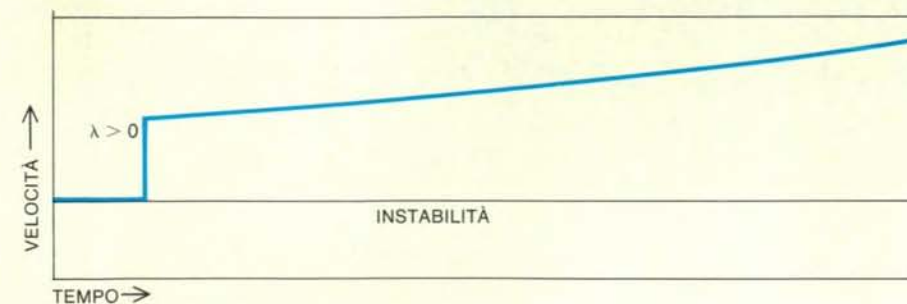
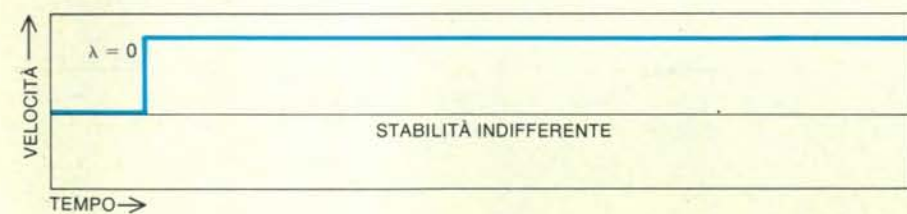
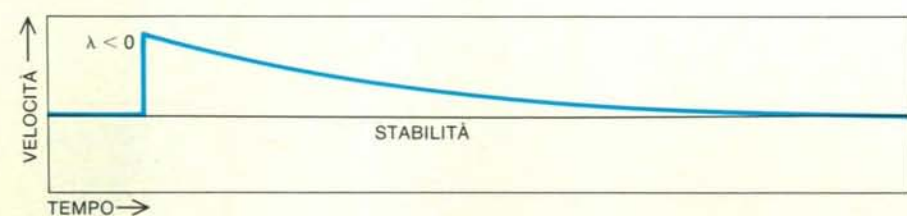
Un altro modello di superficie di potenziale si ricava capovolgendo la boccia emisferica e disponendo accuratamente in equilibrio la biglia sul suo vertice. Anche questo è uno stato di equilibrio, nel senso che tutte le forze agenti sulla biglia sono equilibrate, e in assenza di qualsiasi perturbazione, la biglia potrebbe restare indefinitamente in quiete. In pratica, però, ci sarà sempre qualche influenza esterna (per esempio, una corrente d'aria o un camion in transito) che disturberà questo equilibrio precario se lo sperimentatore aspetta abbastanza a lungo. Dopo una siffatta perturbazione la biglia non ritorna al punto di equilibrio ma se ne allontanerà indefinitamente. Per quanto piccola sia la perturbazione iniziale, la biglia si troverà alla fine a una grande distanza dal centro. La perturbazione è amplificata e quindi lo stato di equilibrio è detto instabile.

C'è una terza possibilità: la biglia può anche essere posta su una superficie piana. In questo caso, se spostiamo la biglia essa non ritorna nella sua posizione originale né continua ad allontanarsi da essa, ma si limita a restare in equilibrio nella sua nuova posizione. Si dice che ogni punto di una superficie di potenziale piana rappresenta uno stato di equilibrio indifferente o marginale.

Se meditiamo un po' di più su questo modello di superficie di potenziale si conclude che la stabilità assoluta di un sistema può essere dimostrata soltanto verificando la sua risposta a tutte le perturbazioni possibili. Per esempio, una biglia all'interno della boccia ritornerà al centro dopo una perturbazione infinitesima o anche dopo una piccola, ma finita. Non ritornerà invece al suo punto di partenza se la perturbazione è sufficientemente grande da scagliarla completamente all'esterno della boccia. Poiché si dovrebbero verificare un numero infinito di possibili perturbazioni, è difficile dimostrare se uno stato di equilibrio è stabile; invece l'instabilità si può dimostrare mettendo in evidenza anche una sola perturbazione che cresce spontaneamente.

L'applicazione di questi principi al problema della convezione è immediata. Si dimostra facilmente che uno strato di fluido immobile riscaldato uniformemente dal sotto è in uno stato di equilibrio, anche se si trascurano lo smorzamento viscoso e la diffusione del calore. Anche se la porzione più leggera del fluido è sovrastata da materiale più denso, in modo da poter diminuire l'energia potenziale gravitazionale scambiando le due posizioni, in assenza di perturbazioni tutte le forze che agiscono su una qualsiasi porzione di fluido sono in equilibrio. La domanda alla quale deve rispondere una teoria della convezione è se l'equilibrio è stabile o instabile o se presenta invece una stabilità indifferente. In altre parole, la teoria deve definire la forma della superficie di potenziale.

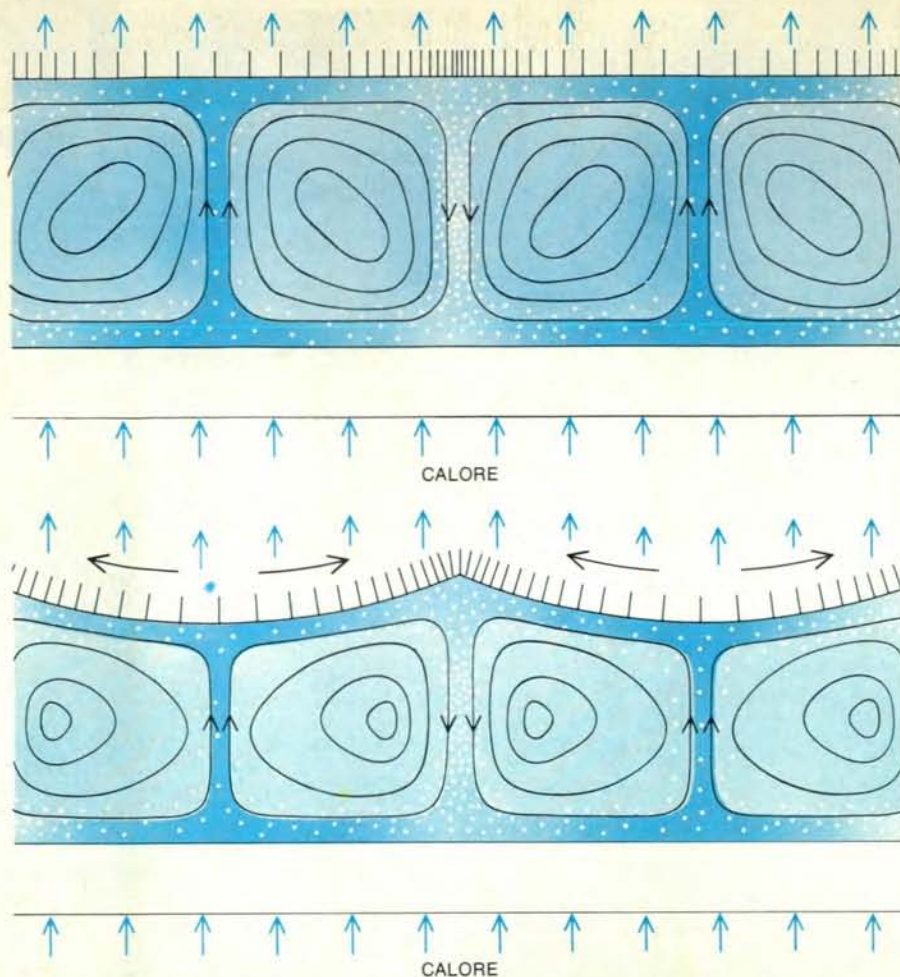
È il valore del numero di Rayleigh a determinare la curvatura della superficie di potenziale. Se il numero di Rayleigh è nullo, per esempio perché sono nulli il gradiente termico e la forza di spinta, lo stato di quiete è ovviamente stabile e la superficie di potenziale ha la concavità verso l'alto, come l'interno di una boccia o di una valle. Per mettere in moto il fluido è necessario aumentarne l'energia. D'altra parte, se il numero di Rayleigh è molto grande, in modo tale che la spinta verso l'alto superi tutti gli effetti dissipativi, il fluido può diminuire la sua energia totale dando luogo a un flusso convettivo. In tal caso qualsiasi perturbazione dell'equilibrio stazionario verrà amplificata. La superficie di potenziale ha la concavità verso il basso come una boccia capovolta.

La continuità impone che esista qualche valore del numero di Rayleigh, intermedio tra questi estremi, in corrispondenza del quale la spinta verso l'alto e le forze dissipative abbiano lo stesso modulo. È questo il numero di Rayleigh critico, che caratterizza la superficie di potenziale piana, di transizione o di stabilità indifferente. Quando il numero di Rayleigh

La stabilità di un sistema fisico si può giudicare dalla sua risposta a una perturbazione arbitraria, quale un piccolo spostamento di una biglia in quiete su una superficie. Se la superficie è concava, la biglia ritorna alla posizione di equilibrio sul fondo della boccia: tale posizione è detta di equilibrio stabile. Su una superficie convessa la biglia può essere tenuta in equilibrio sul vertice, ma qui l'equilibrio è instabile: la più piccola perturbazione è amplificata come la biglia diminuisce la propria energia potenziale rotolando verso il basso. Su una superficie piana la biglia non ritorna nella sua posizione di partenza né si allontana ulteriormente da essa: si parla di equilibrio indifferente. Un sistema può essere stabile di fronte a certe perturbazioni, ma non ad altre, come per una superficie che presenta sia regioni concave sia convesse. In un fluido semplice per innescare la convezione è necessaria una distribuzione instabile di alcune proprietà quali la densità o la tensione superficiale.

aumenta a partire da zero (per esempio, a causa di un crescente gradiente termico) la superficie di potenziale parte concava e gradualmente si appiattisce; in corrispondenza del numero di Rayleigh critico, essa è perfettamente piana, mentre, quando il numero continua ad aumentare, la superficie diventa convessa. L'equilibrio di uno stato di quiete diventa instabile soltanto quando si supera il valore critico. Per l'esperimento modello qui considerato, i calcoli indicano un numero di Rayleigh critico di 1708. In un esperimento di laboratorio, dove il fluido è uno strato di olio di silicone di qualche millimetro di spessore, il numero di Rayleigh critico viene raggiunto quando il gradiente termico è di pochi gradi centigradi.

L'esperimento ipotetico su cui si basa questa spiegazione della teoria di Lord Rayleigh comprende molte ipotesi semplificative, alcune delle quali contrarie ai fatti. La teoria ha comunque avuto un notevole successo nel predire le condizioni necessarie per l'inizio della convezione nei fluidi reali. Per esempio, gli esperimenti di Peter L. Silveston dell'Università della Columbia Britannica e di Ernest L. Koschmieder dell'Università del Texas ad Austin hanno fornito valori del numero di Rayleigh critico di 1700 ± 50, in buon accordo con il valore teorico.

L'equilibrio tra la spinta verso l'alto e le forze dissipative fornisce un criterio per l'innesco dell'instabilità convettiva, ma che cosa si osserva nell'esperimento modello una volta iniziato il flusso? La teoria di Rayleigh fornisce soltanto indicazioni limitate nel rispondere a tale domanda e anche le teorie più complesse che discuteremo in seguito non sono in grado di tener conto di tutti gli aspetti osservati in una circolazione convettiva pienamente avviata. Tuttavia, anche se non si può ricavare matematicamente l'evoluzione del flusso, è almeno possibile una descrizione qualitativa.

In uno strato fluido riscaldato uniformemente dal sotto, il gradiente termico dovrebbe essere indipendente dalla posizione orizzontale e la stessa cosa dovrebbe valere per la spinta verso l'alto risultante. Quando si supera il numero di Ray-

La stabilità di un fluido di fronte alla convezione indotta dalla spinta verso l'alto viene determinata dal valore del numero di Rayleigh. Se il numero è inferiore al valore critico, qualsiasi perturbazione viene smorzata, mentre se il numero supera il valore critico, la perturbazione continua a crescere. Il sistema è più sensibile a perturbazioni con un particolare numero d'onda, o scala di lunghezza, corrispondente al doppio della profondità dello strato fluido. La differenza tra il numero di Rayleigh effettivo e il numero critico determina un parametro λ; nella teoria proposta da Rayleigh la velocità del flusso convettivo dipende da λ in modo esponenziale. Se λ è negativo, la velocità scende a zero; se λ è nullo la velocità resta costante; se λ è positivo, la velocità continua ad aumentare. Dato che λ è un numero complesso che può avere un valore immaginario, il flusso può oscillare, fenomeno detto superstabilità.

Le variazioni di tensione superficiale modificano il modello del trasporto convettivo in un fluido dotato di superficie libera. Il valore della tensione superficiale varia al variare della temperatura, ed è maggiore quando il fluido è più freddo. Qualsiasi differenza di temperatura attraverso la superficie dà quindi origine a un gradiente di tensione superficiale. Lo stato di quiete diventa instabile se il gradiente è abbastanza grande da vincere gli effetti dissipativi della viscosità e della diffusione termica. Il valore della tensione viene in questo caso rappresentato dalla intensità del colore. Il fluido viene trascinato lungo la superficie verso regioni più fredde caratterizzate da maggiore tensione superficiale e viene sostituito da fluido più caldo proveniente dal basso.

leigh critico e l'equilibrio in stato di quiete diventa instabile, il fluido caldo tende a salire verso l'alto dovunque, mentre quello freddo tende a discendere dovunque. Ovviamente i due processi non possono verificarsi contemporaneamente. Questo ostacolo viene evitato dalla divisione spontanea dello strato in un modello di «celle» di convezione, in ognuna delle quali il fluido circola in un'orbita chiusa. Argomentazioni teoriche danno un'indicazione approssimativa della scala favorita dei singoli elementi nel modello di convezione. Queste argomentazioni derivano dalla sensibilità variabile dello stato marginalmente stabile a perturbazioni con differenti scale di lunghezza. Qui dobbiamo stare attenti a non confondere l'ampiezza di una perturbazione, che nell'esperimento modello corrisponde allo spostamento verticale di una porzione di fluido, con la scala della perturbazione, che misura le dimensioni della porzione. Se la teoria di Rayleigh deve fornire risultati significativi, l'ampiezza deve sempre essere infinitesima, mentre la scala può raggiungere le dimensioni massime consentite dall'apparecchiatura.

Si è soliti esprimere la scala di una perturbazione in termini di un numero d'onda, che è il reciproco di una lunghezza. Questa abitudine riflette il fatto che la geometria di una perturbazione è generalmente complessa, in modo tale da non avere una sola dimensione chiaramente definita; la perturbazione può essere però decomposta in uno spettro di modi fondamentali, o frequenze spaziali, proprio come un suono complesso può essere analizzato in una combinazione di note pure. Un numero d'onda rappresenta i contributi di una particolare scala di lunghezza alle fluttuazioni casuali. Numeri d'onda maggiori corrispondono a scale più piccole.

La stabilità dello stato stazionario in quiete è suscettibile di essere sconvolta da perturbazioni con certi numeri d'onda piuttosto che con altri. Si può immaginare un esperimento in cui venga misurato il numero di Rayleigh critico in un fluido le cui fluttuazioni possono essere controllate in modo che siano sempre caratterizzate da un solo numero d'onda. Un siffatto esperimento dovrebbe rivelare che l'instabilità si instaura più rapidamente quando il numero d'onda descrive perturbazioni con una dimensione orizzontale circa doppia della profondità del fluido. Per numeri d'onda maggiori o minori, sono necessarie condizioni più estreme (un maggior numero di Rayleigh) per indurre la convezione. Il valore calcolato 1708 per il numero di Rayleigh critico è quello trovato quando le fluttuazioni hanno le dimensioni più adeguate.

La sensibilità del fluido a perturbazioni su una particolare scala implica che tali perturbazioni saranno amplificate più rapidamente di quasiasi altra quando lo strato diventa instabile. Pertanto il modello osservato quando ha inizio la convezione potrebbe presentare qualche peculiarità in prossimità di tale scala. Non è certo che tali peculiarità persistano una volta pienamente sviluppato il flusso convettivo, ma risulta che ciò avviene, purché il numero di Rayleigh non superi di troppo il valore critico.

Il numero d'onda precisa la scala complessiva del modello ma non la sua forma dettagliata; si potrebbero costruire celle di convezione di molte forme differenti in modo che abbiano lo stesso numero d'onda. Il modello osservato in realtà dipende fortemente dalla geometria dell'apparecchiatura sperimentale; esso non può essere ricavato dai principi primi, ma esistono regole empiriche ben consolidate che fanno previsioni qualitative attendibili.

In esperimenti quale quello ipotetico ora descritto, nel quale le superfici superiore e inferiore sono vincolate da contorni rigidi, l'elemento fondamentale del modello è un «rullo» che ha una lunga forma tubolare. Il fluido caldo sale lungo un bordo del rullo, attraversa la superficie superiore perdendovi calore e ricade poi sul fondo dello strato lungo il bordo opposto. La circolazione trasporta quindi il fluido attraverso il contorno inferiore, dove la sua temperatura aumenta nuovamente. Rulli adiacenti hanno versi di rotazione opposti.

Quando si osserva un rullo in sezione trasversale, la sua forma è all'incirca quadrata: la larghezza è uguale all'altezza e quest'ultima è ovviamente determinata dalla profondità dello strato fluido. In tal modo le proporzioni del rullo sono costanti, ma le sue dimensioni dipendono dalla profondità del fluido. Dato che l'unità ricorrente in questo modello è formata da due rulli ruotanti in senso opposto, la scala del modello è adeguatamente definita dalla larghezza di due rulli. Quindi la scala è uguale al doppio della profondità del fluido, secondo le previsioni dell'analisi con numeri d'onda.

La pianta del modello a celle (il suo aspetto dall'alto) è determinata in larga misura dalla forma dell'apparecchiatura. Dato che nella teoria non entrano direttamente dettagli di questo genere non si può facilmente ricavare la pianta; tuttavia l'osservazione sperimentale fornisce un'ampia base di previsione. Se il recipiente è rettangolare, i rulli tendono ad

allinearsi parallelamente ai lati più corti. La larghezza di ogni rullo e quindi il loro numero sono determinati dalla profondità dello strato. In un recipiente cilindrico i rulli formano anelli concentrici.

Come abbiamo detto sopra, l'analisi della convezione di Lord Rayleigh è stata ispirata principalmente dalle osservazioni sperimentali fatte da Bénard verso il 1900. Oggi sappiamo che la teoria di Rayleigh non è adatta per il meccanismo convettivo studiato da Bénard. Le condizioni sperimentali impiegate da Bénard differivano di poco ma in modo sostanziale da quelle qui descritte e l'importanza delle variazioni appare subito nel modello del flusso convettivo. Nella convezione di Bénard i rulli imposti dalla geometria dell'apparecchiatura possono apparire per breve tempo all'inizio del flusso, ma danno subito luogo a una figura più complessa: un mosaico poligonale della superficie fluida. Inizialmente i poligoni sono alquanto irregolari, avendo da quattro a sette lati, con numero medio di sei. Quando la figura si sviluppa pienamente, diventa uno schieramento quasi perfetto di esagoni regolari, disposti come in un alveare. Il centro di ogni cella esagonale è una regione di fluido caldo in movimento verso l'alto, che si estende sulla superficie superiore e ridiscende sul perimetro, dove le celle adiacenti sono unite.

Negli esperimenti di Bénard, come nell'esperimento ipotetico descritto in precedenza, il fluido forma uno strato sottile riscaldato dal sotto. La differenza essenziale è che lo strato non è confinato tra due contorni rigidi, ma è invece esposto all'aria sulla sua superficie superiore. Essendo la superficie libera, il flusso può essere influenzato dalla tensione superficiale; in effetti, è oggi nota l'influenza predominante della tensione superficiale nella convezione di Bénard, dove è più importante della forza di spinta. Non deve pertanto sorprendere che la teoria di Rayleigh, nella quale si assume esplicitamente che la spinta verso l'alto sia la sola forza agente, sia inadeguata a spiegare una convezione di questo tipo. Le previsioni sono errate perfino per grandezze fondamentali quale il gradiente termico necessario per innescare il flusso convettivo. Una valida teoria alternativa venne introdotta solo nel 1958 da J. R. A. Pearson dell'Imperial College of Science and Technology di Londra.

La tensione superficiale è la forza di coesione il cui effetto complessivo è quello di rendere minima la superficie di un fluido. Per esempio, una goccia libera di liquido tende ad assumere una forma sferica per effetto della tensione superficiale, essendo quella la configurazione di minima area. La tensione si può pensare come un intreccio di elastici tesi in tutte le direzioni sulla superficie libera. Se in un punto qualsiasi le forze esercitate dai vari elastici non si equilibrano, lo strato superficiale migrerà verso la regione di maggiore tensione fino a raggiungere l'equilibrio. Il flusso superficiale viene trasmesso all'intera massa fluida come risultato della sua viscosità di taglio.

La tensione superficiale può agire come forza propulsiva in un flusso convettivo perché la tensione varia con la temperatura: come la densità, la tensione superficiale diminuisce al crescere della temperatura. Qualsiasi gradiente termico instaurato attraverso la superficie liquida sarà accompagnato da un gradiente di tensione superficiale. Le regioni più fredde presenteranno una maggiore tensione superficiale, mentre in quelle più calde la tensione diminuirà. Se il gradiente di tensione superficiale porta a uno squilibrio delle forze, ne risulterà un flusso.

La forma a mosaico della superficie con celle esagonali è un aspetto caratteristico della convezione indotta da un gradiente di tensione superficiale. Quando la tensione è maggiore, la superficie si raggrinza e l'area diminuisce. Sopra al vertice ascendente nel centro di ogni cella la superficie presenta una depressione; il fluido deve risalire prima di discendere dal bordo della cella.

L'innesco dell'instabilità convettiva nel sistema di Bénard può essere analizzato allo stesso modo in cui si analizza l'innesco di un flusso indotto dalla spinta verso l'alto. Supponiamo che una porzione di liquido caldo venga spostata verso l'alto da qualche fluttuazione casuale. Che tale moto sia mantenuto o meno da forze di spinta, esso avrà comunque un effetto sulla superficie dello strato, facendo aumentare lievemente la temperatura e riducendo perciò la tensione superficiale nell'area direttamente al di sopra della fluttuazione. Le forze sulla superficie rimangono comunque in equilibrio perché la superficie circostante esercita una trazione uguale in tutte le direzioni su tale regione. Per poter dare inizio al flusso è necessaria una seconda perturbazione in grado di provocare uno spostamento orizzontale di qualche piccolo elemento della superficie nella zona di minor tensione. Le forze di tensione che agiscono sull'elemento spostato sono allora non equilibrate e, se il gradiente di tensione superficiale è abbastanza grande, lo spostamento sarà amplificato. L'elemento di superficie sarà trascinato nella regione più fredda di maggiore tensione e trascinerà con sé parte della massa fluida. Dagli strati più caldi sottostanti verrà quindi risucchiato verso l'alto più liquido, aumentando i gradienti superficiali di temperatura e di tensione. Nel frattempo il fluido che si è raffreddato durante l'attraversamento della superficie comincerà a scendere e si formerà il modello cellulare.

Come nella convezione indotta dalla spinta verso l'alto, la semplice esistenza di un gradiente termico non garantisce il mantenimento di un flusso convettivo. Il gradiente deve essere abbastanza grande da vincere gli effetti dissipativi della resistenza viscosa e della diffusione del calore. L'equilibrio di tali effetti nel sistema di Bénard viene espresso da un altro rapporto adimensionale, chiamato in questo caso con il nome del ricercatore italiano del XIX secolo C. G. M. Marangoni. La formula è la stessa del numero di Rayleigh con la differenza che la forza di spinta è sostituita dalla forza di tensione superficiale; in altre parole il numero di Marangoni è il rapporto tra il gradiente di tensione superficiale e il prodotto della resistenza viscosa per la diffusione del calore. La convezione di Bénard ha luogo quando il numero di Marangoni supera un valore critico.

Un aspetto caratteristico della convezione guidata da gradienti di tensione superficiale è che essa altera il contorno della superficie. Le regioni di elevata tensione superficiale tendono a raggrinzarsi diminuendo in tal modo la loro superficie totale esposta. Le conseguenze di questo effetto vanno contro l'intuito. Nel centro di una cella di Bénard, dove il fluido sta salendo, la superficie presenta una depressione, mentre ai bordi della cella, dove il fluido sta scendendo, la superficie è rialzata. Le forze gravitazionali si oppongono alla formazione di sporgenze della superficie, dal momento che l'energia potenziale gravitazionale è minima su

Un vertice ascendente di fluido caldo crea un gradiente di tensione superficiale lungo la superficie di separazione di due liquidi. La depressione della superficie indica il punto di minima tensione e il fluido viene trascinato lungo la superficie di separazione verso destra e lasciato lontano da tale punto. La circolazione è indotta dalla trazione superficiale. Un meccanismo analogo agisce in un fluido la cui superficie è esposta all'aria. La fotografia è di H. Linde dell'Istituto centrale di chimica fisica di Berlino. Le particelle plastiche sospese nei liquidi visualizzano il flusso.

una superficie piana. Così le interazioni della gravitazione e della tensione superficiale sono delicate e complesse. Una teoria che comprende sia le forze di spinta sia quelle di tensione superficiale è stata formulata nel 1964 da D. A. Nield dell'Università di Auckland.

Il predominio della tensione superficiale nella convezione di Bénard è stato ora accertato senza alcun dubbio. Una prova indiscutibile è che le caratteristiche celle esagonali di convezione appaiono anche quando lo strato fluido viene riscaldato da sopra anziché da sotto. In tali circostanze il gradiente di densità si oppone al flusso convettivo, cosicché le forze risultanti dalla tensione superficiale lo devono superare. Il flusso convettivo attribuito alla tensione superficiale è stato osservato anche in esperimenti condotti da due missioni spaziali Apollo, nelle quali la gravità e la spinta verso l'alto erano trascurabili.

La teoria di Rayleigh e le teorie basate su di essa forniscono approssimativamente le condizioni necessarie per l'innesco della convezione. Cosa accade una volta iniziato il flusso? Nella descrizione di tali moti convettivi completamente innescati le teorie di questo genere sono molto meno soddisfacenti.

Nella teoria di Rayleigh la velocità del flusso è data da una funzione esponenziale: la velocità è proporzionale al numero di Eulero, e (che ha un valore numerico di circa 2,7) elevato a una certa potenza. L'esponente è uguale al tempo, t (misurato in secondi, per esempio, a partire da un certo istante di riferimento) moltiplicato per un coefficiente, λ, determinato dal numero di Rayleigh. La velocità è perciò proporzionale a $e^{\lambda t}$. Le previsioni della teoria possono essere catalogate osservando l'evoluzione di questa espressione per differenti valori di λ.

Se il numero di Rayleigh è minore del numero critico di Rayleigh, λ è negativo. Con il passar del tempo, perciò, la velocità è data da potenze negative di e progressivamente crescenti; in tali condizioni il valore dell'esponenziale tende a zero al crescere di t. In altre parole, la velocità cade a zero e tutti i moti casuali del fluido vengono smorzati. Quando il numero di Rayleigh coincide esattamente con il valore ciritico, λ è nullo, e quindi l'esponente λt è sempre nullo. Dal momento che qualsiasi numero elevato a potenza nulla è uguale a 1, la perturbazione non è né smorzata né accelerata, ma mantiene indefinitamente il suo valore iniziale.

Entrambe le previsioni sono in accordo con l'intuizione e con l'analisi della stabilità dello strato fluido. Un valore negativo di λ corrisponde allo stato stabile di quiete, mentre il valore $\lambda = 0$ indica una stabilità marginale. L'interpretazione della teoria diventa però problematica quando il numero di Rayleigh è maggiore del valore critico e λ è positivo. Sono proprio le condizioni nelle quali si può instaurare un flusso convettivo.

Quando λ è positivo, la potenza a cui viene elevato e aumenta continuamente nel tempo e l'espressione riproduce il ben noto andamento di una crescita esponenziale. Se λ è uguale a $+1$ e la velocità iniziale è di un centimetro al secondo, dopo un secondo la velocità è aumentata a 2,7 centimetri al secondo e dopo due secondi ha raggiunto 7,4 centimetri al secondo. La velocità cresce illimitatamente, il che conduce subito a previsioni assurde; per esempio, la corrente di convezione prevista raggiunge la velocità della luce in meno di mezzo minuto.

La dipendenza della velocità del flusso dal valore di λ è stata qui esposta in forma semplificata. In generale λ è un numero complesso, un numero cioè che ha una parte reale e una parte immaginaria; quest'ultima comprende come fattore la radice quadrata di -1. Finora abbiamo considerato soltanto la variazione della parte reale di λ. Se la parte immaginaria non è nulla, può nascere un flusso oscillante, condizione detta di superstabilità. Tali correnti oscillanti sono state osservate nei fluidi reali e costituiscono una sottoclasse interessante dei fenomeni convettivi. Nel modello descritto dalla teoria di Rayleigh, la parte immaginaria di λ però scompare e si deve affrontare direttamente il problema della crescita esponenziale.

L'aumento di velocità di un flusso convettivo non può ovviamente continuare molto a lungo sulla curva esponenziale. Per tale ragione le previsioni della teoria di Rayleigh si possono considerare realistiche soltanto quando il numero di Rayleigh è prossimo al valore critico (in modo che λ sia piccolo) o soltanto per un breve intervallo di tempo successivo all'innesco della convezione (in modo che t sia piccolo). La causa fisica di tali limitazioni sta nelle ipotesi semplificative adottate nello sviluppo della teoria. In particolare abbiamo supposto che il gradiente termico fosse costante e quindi non influenzato dalla circolazione convettiva. Questa ipotesi è chiaramente contraria all'evidenza: quando un fluido caldo risale verso la parte superiore più fredda dello strato, la differenza di temperatura tra i contorni superiore e inferiore diminuisce. Diminuisce proporzionalmente la forza di spinta e il flusso pertanto si autolimita. Tuttavia, questo meccanismo di autolimitazione non compare nella matematica della teoria di Rayleigh, dove il gradiente termico resta costante indipendentemente dalla rapidità del flusso e la forza di spinta fornisce un impulso per un'accelerazione che continua indefinitamente.

Constatato il fallimento, può apparire sorprendente che la teoria di Rayleigh porti in tutti i casi a risultati accettabili. Può fare ciò soltanto perché comprende un'altra ipotesi, cioè che la porzione di fluido non riceve mai più di uno spostamento infinitesimo. Se è rispettata tale condizione, l'ipotesi di un gradiente termico costante è del tutto ragionevole. Perfino un movimento della porzione finito ma piccolo può provocare soltanto una piccola perturbazione della distribuzione di temperatura, in modo tale da lasciare ancora approssimativamente valide le previsioni della teoria. Tuttavia, nell'applicare la teoria di Rayleigh a un flusso completamente sviluppato, le ipotesi vengono violate e l'evoluzione del moto prevista finisce in un poco attendibile infinito.

Una teoria accettabile della convezione ben avviata deve in qualche modo scopri-

re il meccanismo di retroazione attraverso il quale il flusso stesso altera la forza che lo guida. Non si conosce alcun metodo pratico per risolvere esattamente il problema, ma esistono alcune approssimazioni che danno risultati migliori della teoria di Rayleigh. Nel 1937 il teorico sovietico L. D. Landau sviluppò la sua teoria in modo da descrivere certi tipi di transizioni di fase, quale l'innesco della magnetizzazione nel ferro ferromagnetico. Con V. L. Ginzburg, un altro fisico sovietico, egli la adattò più tardi alla descrizione della superconduttività nei metalli. Questi fenomeni hanno certi aspetti in comune con la convezione; in particolare, essi richiedono la descrizione simultanea delle fluttuazioni di struttura su molte scale di lunghezza. Quando si adatta la teoria di Landau al problema della convezione, essa comprende in prima approssimazione la teoria di Rayleigh.

Ciò che la teoria deve fornire è un'equazione di moto del fluido, un'equazione che esprime la velocità e l'accelerazione di una qualsiasi porzione di fluido per qualsiasi data combinazione di condizioni esterne imposte. L'equazione di moto si potrebbe costruire direttamente, ma il significato della teoria di Landau è più chiaro se si definisce invece una superficie di potenziale, dalla quale si può ricavare l'equazione di moto. (L'equazione di moto è semplicemente l'equazione che dà la pendenza della superficie di potenziale.)

La superficie di potenziale può essere immaginata come un panorama ondulato, nel quale l'altezza al di sopra o al di sotto di un qualche piano di riferimento rappresenta l'energia relativa del sistema fluido. La tendenza dell'energia ad assumere un valore minimo implica che il punto che rappresenta lo stato del sistema «rotoli verso il basso» appena possibile. Un asse sul piano di riferimento definisce la linea a velocità nulla; spostamenti di questa linea verso destra o verso sinistra indicano una velocità crescente positiva (verso l'alto) o negativa (verso il basso) di qualche porzione del fluido. La posizione sull'asse a velocità nulla è data dalla differenza tra il valore effettivo del numero di Rayleigh e il valore critico; tale differenza si può indicare con ΔR. Il valore di ΔR e della velocità V indicano insieme un punto sul piano di riferimento; l'altezza della superficie di potenziale in tal punto è l'energia del sistema in tale stato.

È necessario discutere in dettaglio l'equazione che descrive la topografia della superficie di potenziale. L'equazione si può scrivere come la somma di una serie infinita di termini, ciascuno dei quali contiene una potenza sempre più alta della velocità del fluido. Il primo termine è quadratico: $-\frac{1}{2}\Delta R V^2$. Nel termine successivo la velocità compare come V^3, nel successivo ancora come V^4 e così via. Ognuna delle potenze crescenti della velocità deve essere preceduta da un coefficiente che misura il suo contributo alla forma della superficie.

Come ci si potrebbe aspettare, è scomodo calcolare la somma dell'intera serie infinita; anche se disponessimo di metodi formali per farlo, non si conoscono i coefficienti di tutti i termini. Però, in generale si prevede che i coefficienti diventino più piccoli al crescere della potenza alla quale è elevata V. C'è quindi qualche speranza di ricavare una precisione ragionevole da una serie troncata, cioè una serie nella quale si trascurano tutti i termini al di là di una certa potenza di V. Se la velocità non è troppo grande, i contributi di tali termini di grado più elevato dovrebbero essere piccoli; in particolare, se V è minore di 1 in un certo sistema naturale di unità, le potenze più alte di V convergeranno a zero e l'approssimazione sarà buona.

La teoria di Rayleigh può essere vista come un tale troncamento della serie infinita, che tiene valido quindi soltanto il primo termine, cioè $-\frac{1}{2}\Delta R V^2$. La superficie descritta da questa espressione quadratica ha due lobi, ognuno dei quali con una sezione trasversale parabolica; un lato è concavo verso l'alto, mentre l'altro è concavo verso il basso. Dall'equazione è ovvio che quando V è nulla (cioè, in tutti i punti dell'asse a velocità nulla), l'energia relativa è nulla. Se il numero di Rayleigh è inferiore al valore critico, e quindi ΔR è trascurabile, la superficie si incurva verso l'alto ai due lati dell'asse e l'energia aumenta tutte le volte che la velocità è maggiore di zero. In altri termini, lo stato di quiete è un minimo della superficie di potenziale, uno stato di equilibrio stabile. Quando il numero di Rayleigh supera il valore critico, e quindi ΔR è positivo, la situazione è esattamente opposta: la superficie cade ai lati dell'asse a velocità nulla, che ora è una linea di equilibrio instabile e di energia massima.

Queste proprietà della superficie sono proprio quelle previste dall'analisi elementare di stabilità esposta sopra e comprendono sia i pregi sia i difetti della teoria di Rayleigh. Nelle immediate vicinanze dell'origine, dove tanto ΔR quanto V sono piccoli, si può dedurre correttamente il comportamento del sistema dalla variazione di curvatura della superficie. Quando ΔR è di poco negativo il fluido si riporta in quiete dopo una qualsiasi piccola perturbazione. Se ΔR è di poco positivo la perturbazione viene amplificata e inizia un flusso convettivo. Quando il numero di Rayleigh ha esattamente il valore ciritico, cioè ΔR è nullo, la superficie è piatta e una fluttuazione casuale di velocità non viene né smorzata né amplificata. Invece, per valori più alti di ΔR e di V, ci si ritrova di fronte a un problema che ormai dovrebbe esserci familiare. La superficie di potenziale pende verso meno infinito con il risultato di far aumentare illimitatamente la velocità.

La superficie di potenziale associata alla teoria di Rayleigh definisce l'energia relativa del fluido per qualsiasi combinazione del numero di Rayleigh e della velocità. La superficie è descritta da un'equazione quadratica nella quale il coefficiente, ΔR, è la differenza tra il valore effettivo e il valore critico del numero di Rayleigh. Fin quando ΔR è negativo, l'asse a velocità nulla rappresenta lo stato di minima energia e può essere mantenuto un flusso convettivo soltanto a condizione di aumentare l'energia. Quando ΔR è positivo, la pendenza della superficie viene invertita e l'asse a velocità nulla definisce uno stato di equilibrio instabile; il fluido può diminuire la propria energia instaurando una circolazione convettiva. Il difetto maggiore della teoria di Rayleigh è che, una volta iniziato il flusso, la velocità continua ad aumentare indefinitamente, previsione non realistica.

Nella teoria di Landau a questo difetto si può porre rimedio mantenendo altri termini della serie infinita. In realtà si ottengono buoni miglioramenti con la sola aggiunta di pochi termini. Per l'esperimento ipotetico dal quale è stata ricavata la teoria di Rayleigh la scelta dei termini è limitata da considerazioni di simmetria. In tale esperimento la natura della circolazione resterebbe immutata se l'apparecchiatura venisse riscaldata dal sopra anziché dal sotto, o se tutte le velocità invertissero direzione. Questa invarianza implica che la superficie di potenziale sia simmetrica rispetto all'asse a velocità nulla; una siffatta superficie simmetrica è descritta da un'equazione che comprende solo potenze pari della velocità (quali V^2, V^4 e così via). Quando l'esponente è un numero pari, V e $-V$ elevati alla stessa potenza danno lo stesso risultato, mentre i risultati sono di segno opposto quando l'esponente è un numero dispari. Quindi tutti i termini in cui V è elevato a una potenza dispari (quali V, V^3 e così via) devono avere coefficienti nulli.

Interessanti risultati si possono avere aggiungendo un altro termine nel potenziale di Landau oltre a quello quadratico, cioè il termine di quarto grado $\frac{1}{4}V^4$. La topografia della superficie di potenziale si determina in tal caso calcolando l'espressione $-\frac{1}{2}\Delta R\ V^2 + \frac{1}{4}V^4$. Quando ΔR è negativo la superficie assomiglia molto alla più semplice superficie quadratica, pur se l'energia aumenta più rapidamente quando la velocità si allontana da zero. Quando ΔR è positivo, la forma della superficie viene alterata in modo significativo. L'energia diminuisce su entrambi i lati dell'asse a velocità nulla, ma la diminuzione non continua indefinitamente. L'energia raggiunge invece un valore minimo per poi risalire quando la velocità aumenta ancora. Sia il valore del minimo di energia sia la velocità alla quale si raggiunge il minimo aumentano con ΔR.

Per un limitato intervallo di numeri di Rayleigh e di velocità questa versione relativamente semplice della teoria di Landau può fornire previsioni realistiche. Come in precedenza, quando il numero di Rayleigh è minore di quello critico, viene smorzata qualsiasi fluttuazione casuale di velocità e lo stato stazionario è quello di minima energia e di equilibrio stabile. Per valori del numero di Rayleigh superiori a quello critico una perturbazione cresce con maggiore rapidità, ma la crescita non continua illimitatamente. Quando il fluido raggiunge una certa velocità finita, determinata dal valore di ΔR e corrispondente al minimo della superficie di potenziale, si ottiene un nuovo equilibrio stabile. Per qualsiasi valore di velocità diverso da questo si ha smorzamento.

La teoria di Landau che comprende termini quadratici e di quarto grado evita alcuni dei più spettacolari fallimenti della teoria di Rayleigh, ma è ancora un'approssimazione ed è valida soltanto se la velocità non è troppo grande. Quando V è grande, le potenze di V elevate forniscono un contributo significativo anche se sono precedute da un coefficiente piccolo; per questo motivo una teoria che trascura tutti i termini di potenza più alta non può rappresentare con precisione la forma della superficie di potenziale lontano dall'asse a velocità nulla. Inoltre, in molti sistemi convettivi, una direzione di flusso è privilegiata sull'altra, in modo da far perdere la simmetria della superficie di potenziale e nell'equazione devono essere inserite le potenze dispari di V.

Sia la teoria di Rayleigh sia quella di Landau sono ricavate da esperimenti ipotetici nei quali è considerato costante il maggior numero di proprietà del fluido. I fluidi reali sono raramente così semplici e le interazioni delle varie proprietà possono diventare assai complesse. Per esempio, nel modello si era supposto che solo la densità variasse in funzione della temperatura. In realtà anche la viscosità e la diffusività termica variano con la temperatura nella maggior parte dei fluidi. Dal momento che queste grandezze entrano nella definizione del numero di Rayleigh, una loro variazione può avere un'influenza importante sull'innesco di un flusso convettivo e sulla sua conseguente evoluzione. Si era anche supposto che il fluido fosse incomprimibile; dato che molti fluidi reali sono comprimibili, la pressione diventa una variabile significativa, che influenza a sua volta la densità e molte altre proprietà. La temperatura e la viscosità sono correlate da una relazione di notevole complessità. In generale la viscosità diminuisce all'aumentare della temperatura, ma nello stesso tempo l'energia dissipata dalla resistenza viscosa appare sotto forma di calore e fa quindi aumentare la temperatura.

Una teoria che tenga conto esplicitamente di tutte le interazioni conosciute delle proprietà del fluido sarebbe troppo scomoda per essere pratica. Nella descrizione di un sistema di un fluido reale più che un modello matematico ciò che si deve cercare è un valido compromesso tra la complessità del fluido e la complessità della teoria. La natura del compromesso necessario si può illustrare con alcuni esempi di convezione nel mondo esterno al laboratorio.

Nell'atmosfera terrestre si osserva la convezione su diverse scale di lunghezza. Il gradiente termico tra i tropici e i poli induce una circolazione globale che può essere decomposta in almeno tre grandi celle convettive in ogni emisfero. Le di-

La teoria di Landau porta a una superficie di potenziale più attendibile aggiungendo alcuni termini addizionali all'equazione che definisce la superficie. Nel caso più semplice qui illustrato, viene aggiunto un termine di quarto grado al termine quadratico della teoria di Rayleigh. Lo stato stazionario di quiete diventa di nuovo instabile quando ΔR diventa maggiore di zero, ma la velocità non continua a crescere indefinitamente. Appare invece un minimo nella superficie di potenziale per una velocità finita. La teoria prevede un nuovo equilibrio alla velocità che rende minima l'energia totale; essa è stata avanzata nel 1937 da L. D. Landau per descrivere certe transizioni di fase in materiali magnetici ed è stata recentemente adattata al problema della convezione.

storsioni di questi modelli causate dalla rotazione della Terra danno origine ai venti alisei dei tropici e ai venti occidentali prevalenti delle zone temperate. Il riscaldamento locale dell'atmosfera nei pressi della superficie terrestre dà origine a flussi convettivi su scala minore, tra i quali quelli della maggior parte delle tempeste. I cumuli, che si formano quando l'aria calda sale e si raffredda e diventa perciò sovrassatura di umidità, caratterizzano spesso lo sconvolgimento convettivo dell'atmosfera.

Un'analisi teorica della convezione atmosferica deve tener conto della grande comprimibilità dell'aria, che dà origine a un gradiente di densità anche quando la temperatura è costante con la quota. Una descrizione accurata della circolazione atmosferica dovrebbe anche comprendere il riscaldamento per compressione dell'aria quando essa scende in una regione di maggiore pressione. Anche la viscosità e altre proprietà dell'aria variano con la pressione e la temperatura e la presenza di vapore acqueo, che emette calore quando condensa, aggiunge un ulteriore livello di complessità. Le nuvole formatesi come risultato di una circolazione convettiva sono anch'esse instabili a un ulteriore moto convettivo: la nuvola si raffredda nella parte superiore per perdita di calore verso lo spazio e si riscalda nella parte inferiore per la radiazione assorbita dal suolo. Se l'entità di tali effetti è sufficientemente grande, all'interno della nuvola può instaurarsi una cella di convezione.

Nonostante queste complicazioni il moto convettivo nell'atmosfera manifesta spesso gli stessi modelli fondamentali osservati in più semplici esperimenti di laboratorio. Le formazioni rettilinee di nubi chiamate strie o code di cavallo sono prodotte da celle del tipo a rullo; le fotografie dei satelliti rivelano occasionalmente schieramenti di celle poligonali che si estendono per migliaia di chilometri quadrati. Tuttavia i risultati degli esperimenti di laboratorio non si possono semplicemente estrapolare alla scala atmosferica. Le celle di convezione in laboratorio sono sempre larghe quanto alte, mentre le celle atmosferiche sono molto più larghe, di un fattore fino a 50. Inoltre il verso di circolazione in esperimenti su piccola scala è sempre lo stesso (con i gas il flusso è diretto verso il basso nel centro di ogni cella), mentre nell'atmosfera si osservano entrambe le direzioni di flusso.

Anche la convezione negli oceani copre un ampio intervallo di scale, da un paio di metri fino alle dimensioni degli stessi bacini oceanici. I più semplici di tali flussi hanno una spiegazione immediata. Dato che alcune lunghezze d'onda della radiazione solare penetrano negli oceani per alcune decine di metri, l'acqua viene riscaldata fino a una notevole profondità. Il raffreddamento, del resto, è dovuto quasi interamente all'evaporazione e alla perdita di calore per conduzione e per irraggiamento verso l'atmosfera, processi che sono praticamente limitati alla superficie. Il calore viene quindi introdotto negli oceani a un livello inferiore a quello al quale viene asportato, e uno strato d'acqua di parecchi metri di spessore può diventare instabile alla convezione.

La comprimibilità dell'acqua di mare è piccola e può avere effetto sul moto convettivo soltanto alle maggiori profondità, ma c'è un'altra variabile che può avere un importante effetto: la salinità. La densità dell'acqua marina varia non soltanto con la temperatura ma anche con la concentrazione dei sali disciolti; la densità aumenta con la salinità. Il risultato è che possono cooperare due importanti fattori per determinare un gradiente di densità. L'interazione di questi fattori può originare nuovi tipi di moto convettivo che non appaiono quando è presente solo un unico gradiente.

Se la temperatura è più alta sul fondo di uno strato mentre la salinità è maggiore nella parte superiore, entrambi i gradienti favoriscono la convezione. Quando i gradienti di temperatura e di salinità agiscono in direzioni opposte, entrano in gioco effetti più delicati. Se l'acqua calda salata si trova al di sopra dell'acqua fredda dolce il gradiente di temperatura favorisce la stabilità, mentre il gradiente di salinità la ostacola. Anche se i due gradienti opposti cooperano per produrre una densità uniforme, la convezione può a volte instaurarsi per ragioni che hanno a che fare con effetti dissipativi che agiscono in modo diverso sui due gradienti. Il gradiente di temperatura viene dissipato principalmente per diffusione del calore, mentre il gradiente di salinità viene dissipato principalmente dalla diffusione molecolare delle molecole di sale e d'acqua. La diffusione del calore è molto più rapida, spesso di un fattore 100. Inizialmente la temperatura e la salinità dei due strati potrebbero essere regolate per conferire a essi densità esattamente uguali. Se una porzione del fluido caldo salato viene poi spostata verso il basso nello strato freddo dolce, essa perderà calore per molto tempo prima che la diffusione molecolare possa ridurre la propria salinità in modo significativo. Il risultato è che essa diventerà più densa e il moto sarà amplificato.

La disposizione opposta, con il liquido freddo dolce posto al di sopra del fluido caldo salato può portare al fenomeno oscillatorio chiamato superstabilità. Una porzione di acqua salina calda salendo di poco si raffredda, ma mantiene la propria concentrazione salina. Il risultato è che essa diventa più densa di quanto lo fosse inizialmente e ridiscende verso lo strato inferiore. In realtà, essa può superare la propria posizione originale continuando a oscillare attorno a essa. Le oscillazioni possono crescere o venire smorzate, a seconda dei valori dei due gradienti.

Uno dei sistemi convettivi più complessi è quello che funziona apparentemente nel mantello terrestre, creando una catena di dorsali sul fondo del mare e facendo migrare i continenti sulla superficie terrestre. Il calore che induce la circolazione non viene liberato sul contorno, ma nel volume del materiale, principalmente per effetto del decadimento di elementi radioattivi. In tali circostanze si forma un gradiente termico perché il calore viene perso dal sistema soltanto alla superficie, in modo che la temperatura aumenta con la profondità. Non c'è dubbio che il gradiente sia abbastanza grande da indurre la convezione, ma le proprietà del sistema sono così complicate e il mantello è talmente inaccessibile alle misure che la forma e le dimensioni del modello di convezione sono estremamente incerte. La viscosità aumenta bruscamente con la profondità e a una certa quota della zona di convezione il materiale subisce evidentemente una transizione da una fase cristallina a un'altra.

Su una scala molto più piccola si può osservare un interessante processo convettivo di considerevole complessità in una pellicola di vernice o di lacca che si sta asciugando. In questo caso la forza responsabile non è la spinta verso l'alto, ma la tensione superficiale, come negli esperimenti di Bénard. Il meccanismo responsabile in ultima analisi del flusso è l'evaporazione del solvente dalla superficie libera della pellicola. Se qualche perturbazione fa aumentare la velocità di evaporazione in una regione, tale regione si raffredda con conseguente aumento della sua tensione superficiale. Inoltre la tensione superficiale propria dei pigmenti o di altre grandi molecole della pellicola è solitamente maggiore della tensione del solvente, e quindi una mancanza di solvente fa aumentare la tensione superficiale indipendentemente dalla temperatura. Il liquido viene risucchiato attraverso la superficie verso zone di elevata tensione superficiale, dove esso scende fino alla base della pellicola e riprende il ciclo. Però, quando la concentrazione del solvente è diminuita, la viscosità aumenta e il numero di Marangoni finisce con lo scendere al di sotto del valore critico. A questo punto la convezione cessa.

Le celle di convezione nelle pellicole di vernice hanno spesso una forma esagonale, o almeno una forma poligonale che si avvicina a quella ideale di esagoni regolari. Il flusso può provocare l'«accumulo» dei pigmenti che si manifesta nella pellicola asciutta come una colorazione irregolare. In alcuni casi il modello tridimensionale delle celle di convezione rimane congelato nella pellicola asciutta. Va notato che il fenomeno non è sempre indesiderabile: la pittura a «martello» acquista il suo aspetto in questo modo.

La generalità del concetto di convezione è suggerita da questi differenti esempi: lo sconvolgimento spontaneo dell'atmosfera e degli oceani terrestri, e la circolazione della vernice in una pellicola spessa pochi decimi di millimetro. Le teorie che descrivono questi moti fluidi richiedono molte ipotesi semplificative se devono essere di impiego pratico e anche così sono ben lontane dall'essere semplici. È perciò importante sottolineare che queste teorie correlate, dominate da pochi numeri adimensionali, possono render conto di fenomeni in scale tanto differenti.

Come nasce la turbolenza

di Christian Vidal e Jean-Claude Roux

Le volute di fumo che si alzano dall'estremità di una sigaretta, i vortici d'aria surriscaldata visibili d'estate sulle strade e anche i gorghi che si formano qua e là nella corrente di un fiume sono esempi di moti turbolenti che tutti conoscono. La caratteristica comune di questi moti è l'aspetto irregolare, disordinato (spesso si dice caotico) della loro evoluzione nel tempo. Perciò essi sembrano imprevedibili e non riproducibili e, di conseguenza, il loro studio appare molto delicato. Nonostante questa difficoltà evidente, quella che oggi va sotto il nome di turbolenza è divenuta un campo di indagini molto fecondo. Tra i diversi fattori che hanno contribuito allo sviluppo di questo tema di ricerca, tre vanno menzionati per la loro particolare importanza. In primo luogo è apparso a poco a poco evidente che i fluidi in movimento non erano gli unici sistemi a mostrare questo carattere disordinato e quasi imprevedibile; al contrario la turbolenza si presentava nell'evoluzione di tutti i sistemi dotati di un numero sufficiente di gradi di libertà (il numero minimo di variabili necessarie a descrivere in modo completo il sistema): in altre parole, ci si trova in presenza di un fenomeno estremamente generale che interessa numerosi domini, dalla fisica dei plasmi (gas ionizzati) al comportamento delle popolazioni più varie.

D'altra parte la soluzione di diversi problemi tecnici esige se non la padronanza almeno la comprensione della genesi della turbolenza; senza entrare in particolari, tra i problemi più attuali possiamo ricordare la fusione controllata, il miglioramento dei mezzi di trasporto (aerei, marittimi o terrestri), le previsioni meteorologiche. Un'impostazione teorica fruttuosa sul piano euristico è stata finalmente resa possibile nel 1971, quando David Ruelle dell'Institut des Hautes Études Scientifiques e Floris Takens hanno introdotto in matematica il concetto di attrattore strano. Così generalità del fenomeno, motivazioni pratiche ed economiche e progressi analitici hanno nel complesso conferito alla turbolenza il ruolo di un problema al quale la ricerca scientifica doveva e poteva ormai dedicarsi.

Turbolenza debole e turbolenza forte

Il primo campo di sperimentazione studiato fu, com'è naturale, quello della fluidodinamica. Ancor oggi, del resto, gran parte degli esperimenti riguarda due fenomeni di trasporto (di materia e/o di calore) appartenenti alla dinamica dei fluidi, il flusso di Couette e la convezione di Rayleigh-Bénard, e quindi si è ben lungi dall'aver esaurito le possibilità di questo campo. L'esposizione dei principali risultati ottenuti in questo settore ci consentirà di illustrare prima di tutto il metodo di studio utilizzato. Un altro procedimento, che si potrebbe definire semisperimentale, fu adottato nel 1963 da Edward Lorenz per descrivere l'evoluzione dell'atmosfera terrestre: si tratta della risoluzione numerica di un sistema di equazioni differenziali non lineari opportunamente scelto. (Forniremo più avanti un esempio di equazioni di questo genere, quando espliciteremo il modello di Lorenz.) Questa risoluzione richiede naturalmente l'impiego del calcolatore; ciò fa capire l'importanza dello sviluppo degli strumenti informatici al quale oggi assistiamo. Poi, poco più di 10 anni fa, si è rivelata particolarmente interessante per lo studio della turbolenza una terza via, quella della dinamica chimica. In effetti, in casi molto particolari, l'evoluzione temporale di miscele di reagenti offre molti vantaggi rispetto alla dinamica dei fluidi: una varietà di comportamenti almeno altrettanto ampia, una maggiore semplicità di allestimento sperimentale e una precisione più spinta nelle misurazioni. È vero che fin dal 1973 Ruelle aveva previsto l'esistenza di una turbolenza chimica, ma per poter seguire con successo questa strada è stato necessario attendere che fossero compiuti progressi sufficienti nella conoscenza e nel controllo di alcune rare reazioni chimiche oscillanti. Comunque, come vedremo, questo è ormai cosa fatta e si è autorizzati a sperare che la dinamica chimica darà un contributo apprezzabile allo studio e alla comprensione della turbolenza.

Prima di procedere si impone tuttavia un'osservazione: gli specialisti sono giunti infatti alla conclusione che occorre distinguere fra due grandi categorie di regimi turbolenti. La cosiddetta turbolenza forte si presenta nei sistemi aventi un numero elevato di gradi di libertà. Viceversa, ed è questo il risultato fondamentale del lavoro di Ruelle e Takens, bastano tre gradi di libertà per generare quella che oggi viene chiamata turbolenza debole, caratteristica dei sistemi aventi un piccolo numero di gradi di libertà (come nell'esempio del modello di Lorenz considerato più avanti). Si intuisce facilmente come i metodi per affrontare questi due tipi di turbolenza siano piuttosto diversi e, in particolare, come lo studio della turbolenza debole possa fare a meno di una descrizione statistica, che invece è indispensabile nello studio della turbolenza forte. Dato che i progressi più recenti riguardano in primo luogo la turbolenza debole e che la turbolenza chimica, almeno per quanto se ne sa oggi, si ricollega a essa, in questo articolo discuteremo soltanto lavori relativi a questo secondo tipo di turbolenza.

Se si considerano i progressi delle conoscenze sulla struttura microscopica della materia, è un tantino sconcertante constatare che la fisica contemporanea non è ancora venuta a capo dei problemi legati al moto dei fluidi. Eppure si tratta di un fenomeno a scala macroscopica che, di conseguenza, si potrebbe credere meno difficile da interpretare del mondo delle particelle elementari. L'antichità di questo campo di studi (le prime esperienze di Jean-Léonard-Marie Poiseuille sono del 1846 e le equazioni generali del moto di un fluido risalgono addirittura all'inizio dell'Ottocento) dimostra che non è affatto così. Sul piano teorico la difficoltà deriva anzitutto dal carattere non lineare delle equazioni differenziali che esprimono la conservazione della massa, della

quantità di moto e dell'energia all'interno di un fluido continuo. In effetti non solo non sappiamo ancora risolvere in modo generale, per via analitica, i sistemi di equazioni di questo tipo, ma se il numero delle variabili è uguale o superiore a tre la matematica non fornisce quasi informazioni sulle condizioni di esistenza e sulle proprietà delle soluzioni. Anche sul versante sperimentale le difficoltà sono molto serie, all'infuori naturalmente del regime di scorrimento regolare, o laminare, che purtroppo non è molto istruttivo. Fra l'altro, per essere confrontabili tra loro, tutti gli esperimenti devono riferirsi rigorosamente alla stessa situazione. Ciò spiega perché le ricerche si siano organizzate e sviluppate intorno ad alcuni problemi tipici e perché solo i metodi di misurazione e di analisi siano diversi.

Un regime dinamico può essere caratterizzato dalle proprietà topologiche dell'insieme di traiettorie che gli è associato nello spazio delle fasi o in uno spazio derivato. Qui l'evoluzione di una reazione chimica, misurata dalla concentrazione X di una specie, è descritta in un sistema di riferimento X, $\dot{X}$, $\ddot{X}$ (derivate prima e seconda di X rispetto al tempo). Si osserva che le traiettorie, a lungo sovrapposte (*nel riquadro*), si avvolgono a volte in un vortice (molto ingrandito) dove compiono un numero variabile di giri. Ciò dimostra che l'evoluzione del sistema chimico dipende, come la traiettoria che la rispecchia, dalle condizioni iniziali imposte. Questa dipendenza dalle condizioni iniziali è una delle proprietà caratteristiche e bizzarre degli oggetti matematici chiamati «attrattori strani».

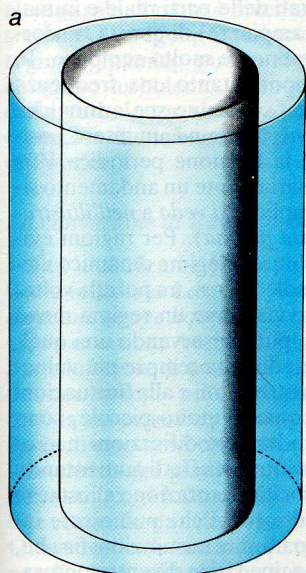
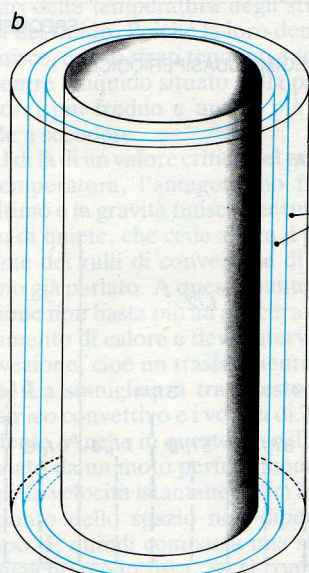
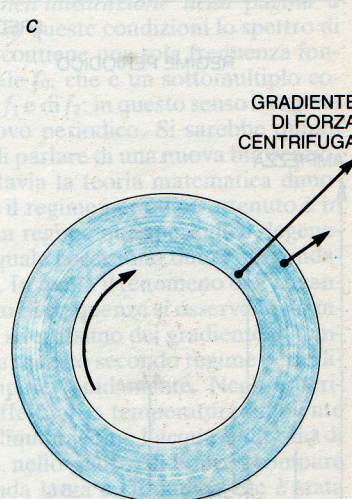
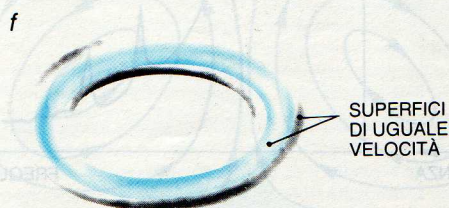
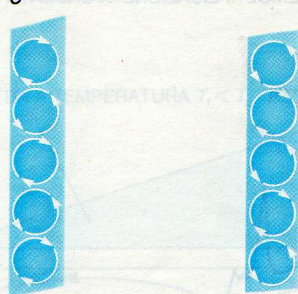
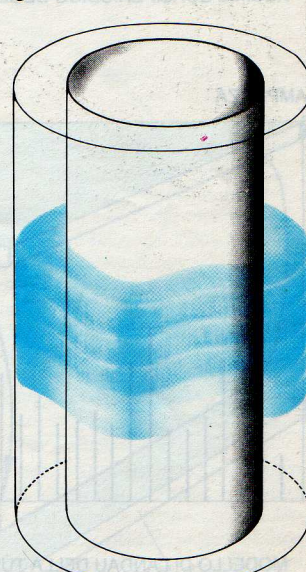
Come abbiamo già accennato, i due fenomeni di gran lunga più studiati in fluidodinamica sono il flusso di Couette e l'instabilità di Rayleigh-Bénard. Nell'ambito di questa presentazione non è certo indispensabile descrivere nei particolari i metodi sperimentali adottati, e quindi non ne parleremo. Non si deve tuttavia dimenticare che, pur senza raggiungere forse il virtuosismo, gli esperimenti implicano il ricorso a tecniche perfezionatissime, in cui hanno una parte notevole l'elettronica e l'informatica. Pertanto i gruppi che lavorano in questo settore non sono molti: particolarmente attivi sono quelli di Günther Ahlers della Bell Telephone, di Pierre Bergé e M. Dubois del Commissariat à l'Énergie Atomique, di Jerry P. Gollub dell'Haverford College, di Albert Libchaber dell'École Normale Supérieure e di Harry L. Swinney dell'Università del Texas ad Austin, per citarne solo alcuni.

Il flusso di Couette

Consideriamo due cilindri coassiali di alcuni centimetri di diametro, tra i quali rimanga uno spazio libero di spessore esiguo (qualche millimetro). Riempiamo questo spazio con un liquido viscoso, per esempio acqua, e facciamo ruotare i due cilindri intorno al loro asse con velocità angolari diverse. La geometria così ottenuta si chiama flusso circolare di Couette, dal nome del fisico francese che ideò il dispositivo. Tra l'ampia gamma di possibilità offerte da questo apparecchio rudimentale, esaminiamo ciò che accade nel caso semplicissimo in cui il cilindro esterno è immobile mentre la velocità di rotazione di quello interno, che all'inizio è nulla, aumenta via via. Il moto del fluido può essere reso visibile aggiungendovi particelle di grafite o di alluminio oppure minutissime scaglie di pesce. Per uno studio quantitativo è necessario misurare una proprietà dinamica, per esempio la velocità istantanea $V(t)$ del fluido in un punto dello spazio.

In un primo tempo, a causa della sua viscosità, il liquido manifesta un moto azimutale. Lo strato di fluido adiacente al cilindro interno è quello che si sposta più velocemente, mentre lo strato a contatto con il cilindro esterno resta quasi immobile. Poiché la forza centrifuga agente su ciascun elemento di fluido aumenta con la sua velocità di rotazione, compare radialmente un gradiente di forza centrifuga (*si veda l'illustrazione nella pagina a fronte*): la precarietà della situazione che si è instaurata è evidente, poiché gli strati che la forza centrifuga tende a spingere maggiormente verso l'esterno sono quelli interni. A un valore critico della velocità di rotazione del cilindro interno, il regime laminare diventa instabile e cede il posto a un altro regime in cui il liquido si organizza in una successione di strati sovrapposti verticalmente l'uno all'altro. Ciascun elemento del fluido è ora animato da un moto avente due componenti fra loro perpendicolari: una, nel piano orizzonta-

le, corrisponde allo spostamento azimutale intorno all'asse dei cilindri; l'altra, nel piano verticale, è associata a un moto circolare che fa passare il fluido dall'interno all'esterno e viceversa. Si è quindi in presenza di una serie di tori orizzontali impilati l'uno sull'altro. Questi tori furono osservati per la prima volta nel 1922 da Sir Geoffrey I. Taylor e da allora sono spesso chiamati vortici di Taylor. Con riferimento alla teoria matematica, si dice che si è prodotta una biforcazione che ha fatto passare il sistema dal regime laminare a quello, diverso sotto il profilo qualitativo, dei vortici di Taylor. Il numero dei tori e il valore critico della velocità di rotazione che li fa apparire dipendono dalla geometria del sistema (altezza e spessore dello strato fluido) e dalla viscosità del liquido, ma si tratta di circostanze particolari che possiamo trascurare.

Il regime dei vortici di Taylor è ancora molto ordinato. Il moto del fluido è perfettamente regolare e, in particolare, la sua velocità $V(t)$ è costante nel tempo in ciascun punto dello spazio: in altri termini, la velocità non dipende dal tempo t bensì dalla posizione dell'elemento di fluido considerato. Fotografando l'insieme a istanti diversi si ottiene sempre la stessa immagine: l'aspetto di questa struttura appare come «congelato». Se si continua ad aumentare la velocità di rotazione, si produce una nuova biforcazione, la struttura cambia e si anima. Le linee di separazione tra i vortici si deformano e assu-

Il flusso circolare di Couette si sviluppa all'interno di uno strato di fluido compreso fra due cilindri rotanti intorno al loro asse. Se il cilindro esterno è mantenuto immobile, quando quello interno comincia a ruotare gli strati del fluido subiscono uno spostamento azimutale tanto più rapido quanto più sono vicini al cilindro. A basse velocità di rotazione del cilindro la velocità del fluido ha solo una componente azimutale (*a*), e le superfici di uguale velocità sono cilindri concentrici (*b*). La forza centrifuga (*c*) tende a spingere verso l'esterno gli strati di fluido più interni. Quando la velocità di rotazione aumenta la situazione diventa instabile. Il fluido si organizza in una serie di tori orizzontali (*d* e, in sezione, *e*), i «vortici di Taylor» (rappresentati solo su parte dell'altezza del cilindro). Le superfici di uguale velocità sono allora dei tori contenuti l'uno dentro l'altro (*f*). Se la velocità di rotazione continua ad aumentare, si ha un altro tipo di instabilità: le superfici di separazione fra i tori divengono ondulate (*g*). Per velocità ancora più alte il regime sarebbe turbolento e la regolarità geometrica apparente del campo di velocità scomparirebbe.

La trasformata di Fourier di una proprietà legata alla dinamica viene utilizzata per caratterizzare i diversi regimi e il modo in cui si instaura la turbolenza. Il regime periodico (*a*) comporta una sola frequenza fondamentale f_1 e alcune sue armoniche. Un regime quasiperiodico (*b*) è associato ad almeno due frequenze fondamentali f_1 e f_2 ed, eventualmente, ad alcune delle loro combinazioni. Questo regime quasiperiodico può degenerare (*c*) quando le due frequenze si «agganciano» su un sottomultiplo comune semplice. In *d* è rappresentata l'evoluzione della trasformata di Fourier quando sopravviene un fenomeno di demoltiplicazione in cascata della frequenza. Secondo la teoria di Lev D. Landau, la turbolenza sarebbe un regime quasiperiodico con un numero infinito di modi fondamentali (*e*). L'esperienza tuttavia dimostra che la turbolenza debole è un regime non periodico la cui trasformata di Fourier contiene una banda larga (*f*) e non righe separate.

mono un andamento sinusoidale. Il loro spostamento azimutale è ben visibile e due fotografie fatte in istanti qualsiasi non sono più, in generale, identiche. La velocità istantanea del fluido $V(t)$ è ormai una funzione periodica del tempo. Questa è la proprietà essenziale del nuovo regime dinamico che è necessario caratterizzare dal punto di vista quantitativo. Un metodo comodo per farlo consiste nel considerare la trasformata di Fourier di $V(t)$. Sotto il profilo fisico ciò equivale a separare la funzione periodica del tempo $V(t)$ nelle sue varie componenti sinusoidali e a calcolarne le ampiezze rispettive. Nel caso qui illustrato lo spettro di potenza (somma dei quadrati delle parti reale e immaginaria delle ampiezze) di questa trasformata è naturalmente molto semplice. In esso compaiono soltanto una frequenza fondamentale f_1 e alcune sue armoniche di frequenza nf_1 (dove n è un intero positivo), perché la funzione periodica $V(t)$ non ha necessariamente un andamento sinusoidale semplice (*si veda* a *nell'illustrazione di questa pagina*). Per ragioni evidenti, questo nuovo regime dinamico viene detto periodico. Così, un po' alla volta, il sistema si avvia verso un regime meno semplice che, pur conservando una qualche regolarità, diventa sempre più vulnerabile alle perturbazioni e alle fluttuazioni le quali, per quanto molto piccole, sono in grado di produrre modificazioni macroscopiche del moto. In effetti aumentando ancora la velocità di rotazione s'instaura una sorta di disordine che può essere individuato solo tramite misurazioni fisiche, poiché l'immagine visiva diventa confusa. In primo luogo si nota un cambiamento qualitativo dello spettro di Fourier di $V(t)$: una terza biforcazione porta a uno spettro contenente non più una ma due frequenze fondamentali, f_1 e f_2, nessuna delle quali è multipla dell'altra (*si veda* b *nell'illustrazione di questa pagina*). Lo spettro contiene allora parecchie righe corrispondenti a combinazioni lineari del tipo $pf_1 + qf_2$ (dove p e q sono numeri interi). Il regime dinamico non è più periodico, ma il suo spettro di Fourier è ancora formato da righe ben separate: si parla di regime quasiperiodico. Dal punto di vista della fisica del moto, la comparsa della seconda frequenza fondamentale f_2 dimostra che un modo di oscillazione supplementare si sovrappone al modo f_1 già osservato. Purtroppo è difficile vedere a quale spostamento sia associato questo secondo modo di oscillazione, perché le frequenze f_1 e f_2 non stanno in un rapporto semplice (di solito si dice che non sono commensurabili) e quindi il moto risultante ha un aspetto molto complesso. Soltanto nel 1980 il gruppo di Swinney negli Stati Uniti e poi quello di G. Cognet a Nancy sono riusciti a identificare questo secondo modo, che corrisponde a una variazione continua dell'ampiezza delle linee sinusoidali che separano due vortici adiacenti. Dunque una situazione che ai nostri occhi appare piuttosto confusa può essere il riflesso di una struttura profonda in defini-

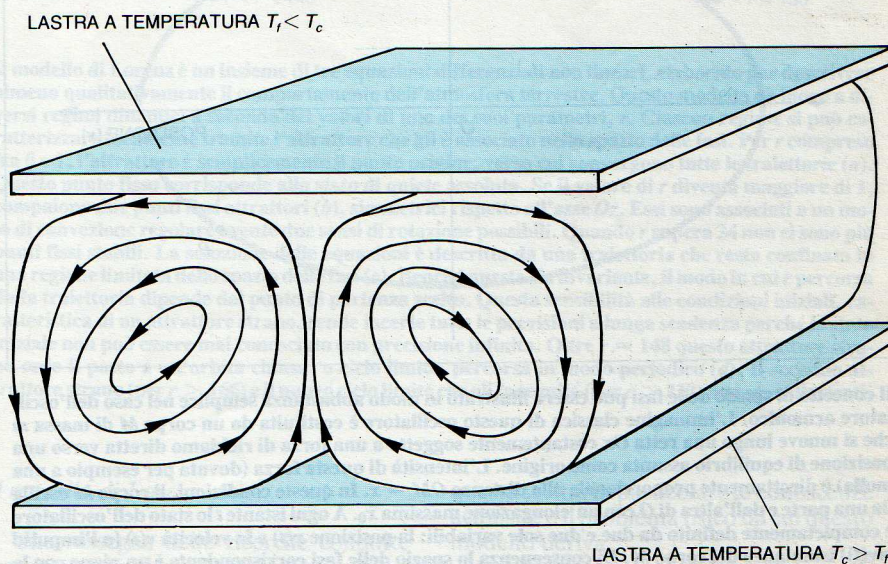
tiva abbastanza semplice. Si tratta di un punto importante che merita di essere sottolineato. Infine, dopo una quarta e ultima biforcazione, nello spettro di frequenza compare una banda larga la cui ampiezza ed estensione aumentano via via (*si veda* f *nell'illustrazione della pagina a fronte*). Allo stesso tempo le righe, presenti fino a quel momento, diminuiscono d'intensità o addirittura scompaiono più o meno completamente. Al contrario delle righe ben separate che si osservavano prima, la presenza di questa banda larga indica che l'energia meccanica si trova ormai distribuita in tutta una gamma di frequenze di cui nessuna è privilegiata. In altre parole, il moto è diventato molto disordinato, irregolare, in breve caotico: il regime dinamico può a ragione essere chiamato turbolento. Dunque la nascita della turbolenza si trova associata alla comparsa di una banda larga nello spettro di Fourier.

La convezione di Rayleigh-Bénard

Nel 1900 Henri Bénard pubblicò un articolo intitolato *Les tourbillons cellulaires dans une nappe liquide*, nel quale riferiva un'osservazione curiosa. Riscaldando la parte inferiore di un sottile strato di liquido (per esempio olio di balena o spermaceti) la cui superficie fosse a contatto con l'aria libera, notò che, oltre un certo valore critico del gradiente di temperatura nello strato, si formava una struttura stabile costituita da celle esagonali contigue. Queste celle di Bénard, il cui asse è perpendicolare allo strato di liquido, sono dovute a un moto convettivo del fluido, che sale alla superficie nel centro di ciascun esagono e ridiscende lungo i suoi margini. L'interpretazione teorica di questo effetto fu intrapresa solo più tardi, nel 1916, in alcuni lavori di Lord Rayleigh, per cui si dà spesso al fenomeno il nome di instabilità o di convezione di Rayleigh-Bénard. Si nota subito ciò che distingue questo problema di trasporto del calore dal flusso circolare di Couette, che abbiamo descritto sopra e nel quale si ha soltanto trasporto di materia. Tuttavia anche se la geometria, il gradiente e le equazioni non sono gli stessi nei due casi, si pone sempre lo stesso interrogativo, cioè quale sia la natura dei moti che nascono spontaneamente in un fluido sotto l'effetto di una sollecitazione esterna. Le analogie, come si vedrà, non mancano.

Per condurre a buon fine uno studio quantitativo bisogna adattare un po' il dispositivo originale di Bénard racchiudendo il fluido tra due lastre, in modo da mantenere e controllare con sufficiente precisione il gradiente di temperatura. Ciò comporta conseguenze apprezzabili, poiché la struttura convettiva cambia. In queste condizioni al posto delle celle di Bénard si osservano rulli il cui asse orizzontale è allineato perpendicolarmente alla dimensione laterale maggiore del dispositivo. Questo esempio illustra senza ambiguità la necessità assoluta di precisare bene tutti i dati sperimentali. In ciò che segue descriveremo i risultati che si ottengono quando la cella di misurazione ha un rapporto di formato (ossia il rapporto tra dimensione laterale e spessore) da uno a tre circa, in modo che i rulli di convezione siano in numero molto piccolo, due o al massimo tre (*si veda l'illustrazione in questa pagina*).

L'aumento progressivo del gradiente di temperatura induce una successione di biforcazioni. A gradienti molto bassi il liquido resta in quiete perché il trasferimento di calore dal basso verso l'alto è assicurato dalla conduzione. Il riscaldamento della lastra inferiore provoca tuttavia un aumento della temperatura negli strati più bassi del fluido. Poiché la loro densità diminuisce, questi strati tendono a sollevarsi mentre il liquido situato nella parte superiore, più freddo e quindi più denso, tende a scendere.

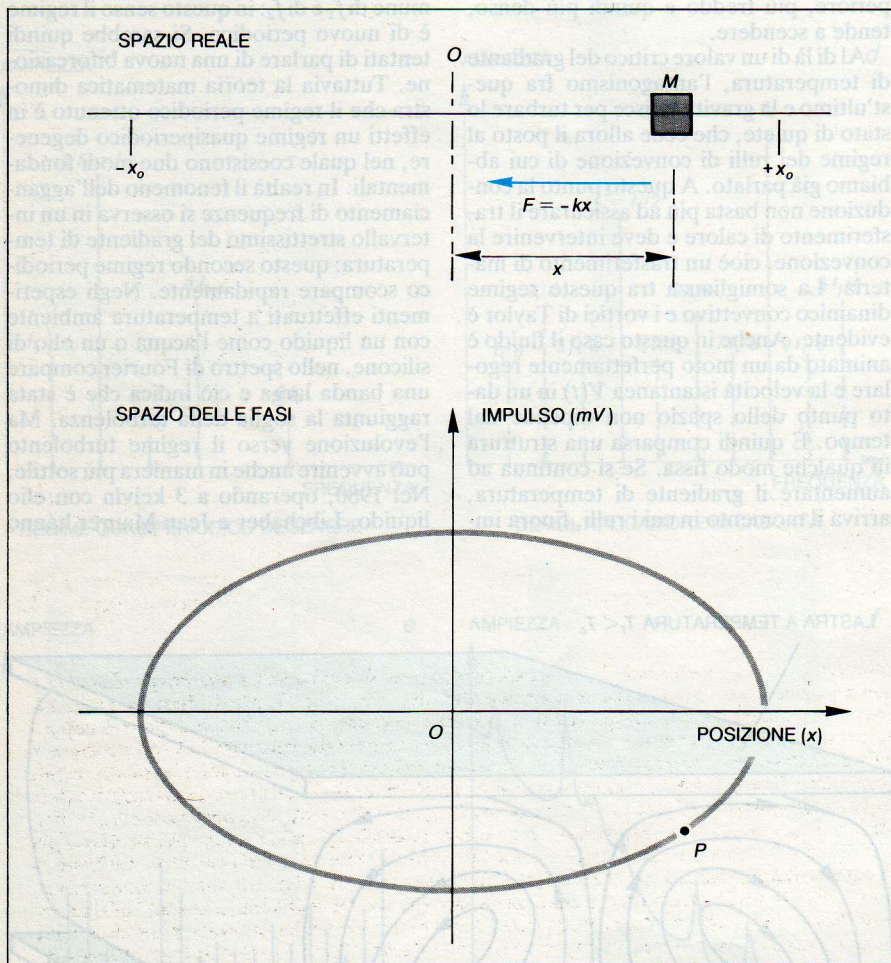
Al di là di un valore critico del gradiente di temperatura, l'antagonismo fra quest'ultimo e la gravità finisce per turbare lo stato di quiete, che cede allora il posto al regime dei rulli di convezione di cui abbiamo già parlato. A questo punto la conduzione non basta più ad assicurare il trasferimento di calore e deve intervenire la convezione, cioè un trasferimento di materia. La somiglianza tra questo regime dinamico convettivo e i vortici di Taylor è evidente. Anche in questo caso il fluido è animato da un moto perfettamente regolare e la velocità istantanea $V(t)$ in un dato punto dello spazio non dipende dal tempo. È quindi comparsa una struttura in qualche modo fissa. Se si continua ad aumentare il gradiente di temperatura, arriva il momento in cui i rulli, finora immobili, si deformano in maniera periodica. Lo spettro di Fourier di $V(t)$ possiede naturalmente una frequenza f_1 ben identificata: il regime è periodico. Un'altra biforcazione porta poi a un regime quasiperiodico con due frequenze fondamentali, f_1 e f_2. Ritroviamo così esattamente gli stadi toccati dal flusso circolare di Couette sotto l'effetto di una velocità di rotazione crescente. A partire da questo punto, tuttavia, la convezione di Rayleigh-Bénard assume un carattere un po' particolare. Si è in effetti osservato un «agganciamento» delle frequenze f_1 e f_2 che, a un dato momento, divengono commensurabili, ossia stanno fra loro in un rapporto semplice (*si veda* c *nell'illustrazione della pagina a fronte*). In queste condizioni lo spettro di Fourier contiene una sola frequenza fondamentale f_0, che è un sottomultiplo comune di f_1 e di f_2: in questo senso il regime è di nuovo periodico. Si sarebbe quindi tentati di parlare di una nuova biforcazione. Tuttavia la teoria matematica dimostra che il regime periodico ottenuto è in effetti un regime quasiperiodico degenere, nel quale coesistono due modi fondamentali. In realtà il fenomeno dell'agganciamento di frequenze si osserva in un intervallo strettissimo del gradiente di temperatura: questo secondo regime periodico scompare rapidamente. Negli esperimenti effettuati a temperatura ambiente con un liquido come l'acqua o un olio di silicone, nello spettro di Fourier compare una banda larga e ciò indica che è stata raggiunta la soglia della turbolenza. Ma l'evoluzione verso il regime turbolento può avvenire anche in maniera più sottile. Nel 1980, operando a 3 kelvin con elio liquido, Libchaber e Jean Maurer hanno

L'instabilità di Rayleigh-Bénard compare in uno strato di fluido posto tra due lastre orizzontali, di cui quella inferiore ha una temperatura più elevata. Oltre un certo gradiente di temperatura ($T_c - T_f$) il trasferimento di calore dal basso verso l'alto non può più essere effettuato dal solo meccanismo di conduzione attraverso lo strato di liquido. Nasce spontaneamente un moto di convezione che, nella geometria qui illustrata, assume la forma di rulli orizzontali. Quando il gradiente di temperatura aumenta ulteriormente, questi rulli si deformano periodicamente. Per valori abbastanza grandi della differenza $T_c - T_f$ si ottiene un regime convettivo turbolento.

messo in evidenza, tramite esperimenti notevolissimi di cui si possono immaginare tutte le difficoltà, un fenomeno di demoltiplicazione in cascata della frequenza a partire da uno stato a due frequenze, cioè $f_0 = f_1 = 2f_2$. Ciò significa che nello spettro di Fourier compaiono successivamente dei sottomultipli, nella fattispecie $f_1/4$, poi $f_1/8$ e $f_1/16$ e alcune delle loro armoniche di ordine dispari (*si veda* d *nell'illustrazione di pagina 144*). È la prima e finora unica volta in cui si sia riusciti a osservare questo particolarissimo comportamento, peraltro previsto dalla teoria.

In ultima analisi questi esperimenti sulla dinamica dei fluidi ci insegnano già molte cose. In primo luogo abbiamo visto che la turbolenza debole s'instaura in poche tappe: possono bastare quattro biforcazioni. Si tratta di un risultato molto importante, che conferma la teoria più recente, quella di Ruelle e Takens, e contraddice le più vecchie idee di Lev D. Landau. Inoltre si osserva che un'apparenza di disordine e di irregolarità può benissimo nascondere un'organizzazione in profondità. Non bisogna certo estendere senza cautele queste ipotesi ad altri campi, poiché nulla ne garantisce la validità al di fuori del dominio sperimentale esaminato. Tuttavia si tratta di un'osservazione molto interessante, che conviene tener presente quando ci si accosta a qualunque problema di dinamica. Infine, e forse soprattutto, è evidente che bisogna perseverare su una strada che a priori non prometteva di essere tanto feconda.

Le due impostazioni teoriche

Per meglio rendersi conto della posta in gioco, è utile ricordare in forma molto qualitativa le due teorie rivali, quella proposta nel 1944 da Landau e quella di Ruelle e Takens. A questo scopo esamineremo brevemente le diverse soluzioni che può avere un sistema di equazioni differenziali rispetto al tempo. In generale si sa che per tutti i problemi di questo genere è comodo ricorrere a una rappresentazione di uno spazio particolare, noto con il nome di spazio delle fasi (*si veda l'illustrazione in questa pagina*). Ciascuna dimensione di questo spazio corrisponde a una variabile indipendente o grado di libertà del sistema.

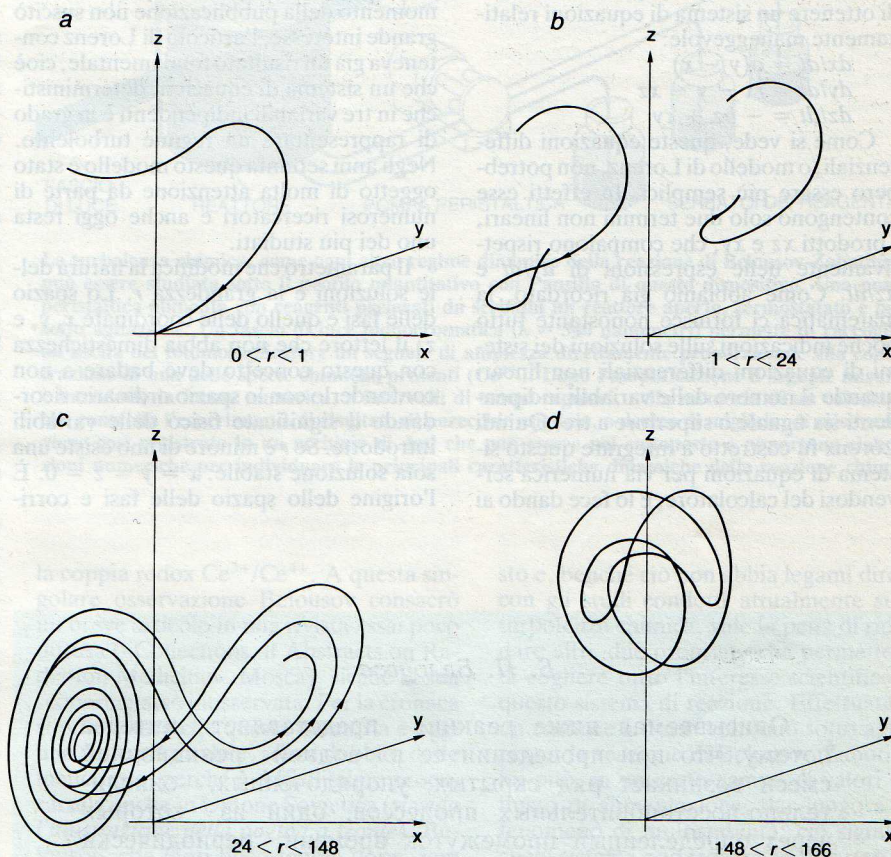
Una soluzione istantanea delle equazioni è rappresentata in questo spazio da un punto le cui coordinate siano i valori delle variabili nell'istante considerato. Se c'è evoluzione, il punto rappresentativo del sistema si sposta, percorrendo una traiettoria nello spazio delle fasi. Questa rappresentazione, che a tutta prima può sembrare una complicazione inutile, in realtà semplifica parecchio la descrizione. Servendoci di essa, consideriamo ora le diverse situazioni alle quali può condurre un sistema di equazioni le cui soluzioni dipendano da un certo parametro. A seconda dei valori di quest'ultimo, il sistema può avere a priori soluzioni molto diverse sotto il profilo qualitativo: può esistere per esempio una soluzione stazionaria, cioè indipendente dal tempo. Nello spazio delle fasi, questa soluzione sarà rappresentata da un punto fisso: l'evoluzione del sistema a partire da diverse condizioni iniziali che non corrispondono a questa soluzione stazionaria sarà rappresentata da un insieme di traiettorie nello spazio delle fasi che convergono verso il punto fisso (purché naturalmente questa soluzione sia stabile e unica, cosa che qui ammettiamo implicitamente). Quindi a ogni soluzione stazionaria stabile è associato un punto attrattore fisso nello spazio delle fasi. Per altri valori del parametro la soluzione può benissimo essere periodica e non più stazionaria: in questo caso la traiettoria è un'orbita che il punto rappresentativo percorre con una certa frequenza f_1. In seguito ai lavori di Henri Poincaré questo attrattore periodico è conosciuto in matematica con il nome di ciclo limite. Naturalmente nulla ci impedisce di considerare anche altre soluzioni, per esempio una funzione periodica che faccia intervenire non una sola frequenza f_1, bensì due frequenze indipendenti f_1 e f_2: l'attrattore allora ha forma toroidale e più in generale è un toro di «dimensione k» se le frequenze indipendenti sono non due ma k. In questa ipotesi la soluzione presenta, nel

Il concetto di spazio delle fasi può essere illustrato in modo abbastanza semplice nel caso dell'oscillatore armonico. L'immagine classica di questo oscillatore è costituita da un corpo M di massa m che si muove lungo una retta Ox costantemente soggetto a una forza di richiamo diretta verso una posizione di equilibrio assunta come origine. L'intensità di questa forza (dovuta per esempio a una molla) è direttamente proporzionale alla distanza $OM = x$. In queste condizioni, il corpo M oscilla da una parte e dall'altra di O con un'elongazione massima x_0. A ogni istante t lo stato dell'oscillatore è completamente definito da due e due sole variabili: la posizione $x(t)$ e la velocità $v(t)$ [o l'impulso $mv(t)$] istantanee del corpo M. Di conseguenza lo spazio delle fasi corrispondente è un piano con le coordinate posizione e impulso: a ogni stato dell'oscillatore è associato un punto P di questo piano. Se l'oscillatore è a riposo nella posizione di equilibrio O, il suo stato è rappresentato dal punto origine del sistema di coordinate, poiché $x = 0$ e $mv = 0$. Se il corpo oscilla liberamente da una parte e dall'altra di O, il punto rappresentativo P si sposta nel piano generando una traiettoria nello spazio delle fasi. La teoria dimostra che si tratta di un'ellisse con centro nell'origine, purché l'oscillatore sia isolato e non vi sia attrito. Così una curva, cioè un'entità geometrica, può fornire una descrizione completa del moto dell'oscillatore. Il ricorso a una rappresentazione di questo tipo si rivela d'interesse decisivo quando il sistema studiato o la sua dinamica sono relativamente complessi.

corso del tempo, un aspetto molto irregolare che induce a definirla quasiperiodica, perché la molteplicità dei modi di oscillazione non permette di distinguerne alcuno con facilità. Nonostante il suo aspetto disordinato, tanto più cospicuo quanto maggiore è il numero k delle frequenze, questa soluzione possiede tuttavia un chiaro ordine soggiacente. Si vede in particolare che partendo da due condizioni iniziali molto vicine, cioè da due punti dell'attrattore tra loro vicini, le traiettorie non si allontanano di molto, sicché in fin dei conti l'apparente disordine è riproducibile! Esiste però anche un tipo di soluzione, detta non periodica, il cui comportamento è del tutto opposto: due traiettorie uscenti da punti vicini dell'attrattore finiscono sempre per allontanarsi l'una dall'altra. Questa volta il disordine osservato non è soltanto apparente: non è più riproducibile e il risultato ottenuto dipende strettamente dalla definizione delle condizioni iniziali. Si dice che il sistema di equazioni presenta una dipendenza sensibile dalle condizioni iniziali. L'attrattore corrispondente viene detto strano perché, anche se ogni traiettoria vi resta confinata, nulla consente di sapere in anticipo quale cammino essa seguirà.

Per riassumere quanto detto, una profonda differenza separa la soluzione quasiperiodica dalla soluzione non periodica. Qualora le condizioni iniziali siano note con precisione adeguata, il destino della prima può essere previsto; per la seconda questa previsione richiederebbe invece una precisione infinita, che per definizione è inattingibile.

Sono stati Ruelle e Takens a stabilire con grande chiarezza questa distinzione capitale. Essi hanno anche dimostrato che la stabilità di un attrattore strano è molto maggiore di quella di un toro di dimensione k. Infatti, se k è uguale a tre o maggiore, basta una leggera perturbazione delle equazioni per fare scomparire il toro, mentre l'attrattore strano continua a esistere.

Sulla base di questi risultati, Ruelle e Takens hanno avanzato l'ipotesi che la turbolenza, manifestazione tipica della dinamica irregolare, dovesse essere descritta da attrattori strani, cioè che fosse in sostanza un fenomeno non periodico (*si veda* d *nell'illustrazione di pagina 144*). In questo modo si ponevano in antitesi con la proposta di Landau, secondo il quale la turbolenza risulterebbe dalla sovrapposizione di un numero, al limite infinito, di modi fondamentali, sarebbe cioè un fenomeno quasiperiodico generalizzato (*si veda* e *nell'illustrazione di pagina 144*). Lo studio sperimentale del flusso circolare di Couette e della convezione di Rayleigh-Bénard con basso rapporto di formato conferma senza possibilità di equivoci la validità dell'impostazione proposta da Ruelle e Takens.

Nella stessa direzione vanno numerosi altri lavori, tra i quali figura in posizione preminente la soluzione al calcolatore di certi sistemi di equazioni differenziali.

Il modello di Lorenz è un insieme di tre equazioni differenziali non lineari, elaborato per descrivere almeno qualitativamente il comportamento dell'atmosfera terrestre. Questo modello dà luogo a diversi regimi dinamici a seconda dei valori di uno dei suoi parametri, r. Ciascun regime si può caratterizzare facilmente tramite l'attrattore che gli è associato nello spazio delle fasi. Per r compreso fra 0 e 1, l'attrattore è semplicemente il punto origine, verso cui convergono tutte le traiettorie (*a*). Questo punto fisso corrisponde allo stato di quiete assoluta. Se il valore di r diventa maggiore di 1, compaiono due punti fissi attrattori (*b*), simmetrici rispetto all'asse Oz. Essi sono associati a un moto di convezione regolare avente due sensi di rotazione possibili. Quando r supera 24 non ci sono più punti fissi stabili. La soluzione delle equazioni è descritta da una traiettoria che resta confinata in una regione limitata dello spazio delle fasi (*c*). Benché questa sia invariante, il modo in cui è percorsa dalla traiettoria dipende dal punto di partenza scelto. Questa sensibilità alle condizioni iniziali, caratteristica di un attrattore strano, rende incerte tutte le previsioni a lunga scadenza perché lo stato iniziale non può essere mai conosciuto con precisione infinita. Oltre $r = 148$ questo attrattore strano cede il posto a un'orbita chiusa, o ciclo limite, percorsa in modo periodico (*d*). Il secondo attrattore strano (per $r > 166$) e il nuovo ciclo limite che gli fa seguito (per $r > 233$) non sono illustrati.

Il modello di Lorenz

Una rassegna delle ricerche condotte tramite la simulazione numerica supererebbe ampiamente i limiti di questo articolo. Di conseguenza ci limiteremo a illustrare l'interesse di questo metodo presentando le informazioni ottenute dallo studio di un particolare sistema di equazioni, che va sotto il nome di modello di Lorenz. Sono molte le ragioni a favore di questa scelta, in particolare le conoscenze acquisite e il problema fisico da cui questo modello deriva.

Lorenz, professore al Dipartimento di meteorologia del Massachusetts Institute of Technology, pubblicò nel 1963 un articolo, oggi celebre, dal titolo ***Deterministic Nonperiodic Flow***, nel quale si proponeva di spiegare il comportamento caotico dell'atmosfera terrestre considerandolo come uno strato d'aria sede di un andamen-

to convettivo di tipo Rayleigh-Bénard. Lorenz scrisse le equazioni del moto prendendo in considerazione tre variabili: la ampiezza $x(t)$ del moto di convezione, la differenza di temperatura $y(t)$ tra correnti ascendenti e discendenti e lo scarto di temperatura $z(t)$ rispetto a un profilo lineare. Per effettuare l'analisi introdusse degli sviluppi in serie di cui conservò soltanto i termini di ordine più basso al fine di ottenere un sistema di equazioni relativamente maneggevole:

$$dx/dt = \sigma(y - x)$$
$$dy/dt = rx - y - xz$$
$$dz/dt = -bz + xy$$

Come si vede, queste equazioni differenziali, o modello di Lorenz, non potrebbero essere più semplici. In effetti esse contengono solo due termini non lineari, i prodotti xz e xy, che compaiono rispettivamente nelle espressioni di dy/dt e dz/dt. Come abbiamo già ricordato, la matematica ci fornisce nonostante tutto poche indicazioni sulle soluzioni dei sistemi di equazioni differenziali non lineari quando il numero delle variabili indipendenti sia uguale o superiore a tre. Quindi Lorenz fu costretto a integrare questo sistema di equazioni per via numerica servendosi del calcolatore, e lo fece dando ai parametri σ e b un valore fisso (rispettivamente 10 e 8/3) e facendo variare progressivamente il terzo parametro r. Benché derivi da una semplificazione troppo drastica perché non sia lecito dubitare del suo carattere realistico, il modello di Lorenz si è tuttavia rivelato di una ricchezza straordinaria, e ciò a dispetto della sua innegabile semplicità derivante da una non linearità poco pronunciata. Anche se al momento della pubblicazione non suscitò grande interesse, l'articolo di Lorenz conteneva già un risultato fondamentale, cioè che un sistema di equazioni deterministiche in tre variabili indipendenti è in grado di rappresentare un regime turbolento. Negli anni settanta questo modello è stato oggetto di molta attenzione da parte di numerosi ricercatori e anche oggi resta uno dei più studiati.

Il parametro che modifica la natura delle soluzioni è la grandezza r. Lo spazio delle fasi è quello delle coordinate x, y, e z; il lettore che non abbia dimestichezza con questo concetto deve badare a non confonderlo con lo spazio ordinario ricordando il significato fisico delle variabili introdotte. Se r è minore di uno esiste una sola soluzione stabile: $x = y = z = 0$. È l'origine dello spazio delle fasi e corrisponde a uno stato uniforme dove la convezione è assente. A partire da qualunque situazione iniziale, cioè da qualunque punto di coordinate $[x(t_0), y(t_0), z(t_0)]$, ogni traiettoria nello spazio delle fasi termina nell'origine. In altri termini, anche se esistono all'inizio, i moti di convezione prima o poi scompaiono. Quando r è compreso tra 1 e 24, vi sono due soluzioni stazionarie stabili, rappresentate nello spazio delle fasi da due punti verso i quali convergono tutte le traiettorie (*si veda l'illustrazione a pagina 147*). Questi punti sono attrattori, così come è un attrattore l'origine nel caso in cui r sia minore di uno. Sotto il profilo fisico ciò corrisponde alla comparsa dei rulli di convezione di Rayleigh-Bénard e ciascun punto è associato a uno dei due sensi di rotazione possibili. Si noti che la configurazione dinamica fissa, in cui il moto del fluido è perfettamente regolare e la velocità istantanea in un dato punto non dipende dal tempo, è rappresentata da un punto fisso dello spazio delle fasi, talché l'informazione principale è ben conservata. Se fino a questo momento il modello di Lorenz non ci ha rivelato nulla di particolarmente interessante, le cose cambiano quando r diventa più grande di 24. Oltre questo valore si nota infatti la scomparsa di ogni soluzione stazionaria stabile. Le traiettorie sono attratte in una regione dello spazio delle fasi, all'interno della quale restano completamente confinate. Inoltre, e soprattutto, il loro percorso cambia a seconda delle condizioni iniziali scelte, rivelando la sensibilità alle condizioni iniziali tipica di un attrattore strano. Nel 1963 Lorenz non poteva servirsi di questo concetto, ma la simulazione numerica al calcolatore gli aveva ugualmente consentito di giungere alla conclusione che non è possibile effettuare previsioni meteorologiche a lungo termine a causa della sensibilità del risultato alle condizioni iniziali scelte per compiere il calcolo. Oggi, nonostante i potenti mezzi impiegati per determinare lo stato dell'atmosfera, il limite oltre il quale la previsione non può essere considerata valida viene valutato intorno ai 15 giorni. Questo fenomeno che, ripetiamo, deriva dalla natura intrinseca del problema, viene talvolta indicato con il nome di «effetto farfalla»: anche ammettendo di poter descrivere lo stato dell'atmosfera con un insieme di equazioni deterministiche di cui si conoscono alla perfezione tutti i parametri, l'incertezza sulle condizioni iniziali introdotta dal battito dell'ala di una farfalla basterebbe a falsare qualunque previsione a lunga scadenza.

Il dominio di esistenza di questo attrattore strano è molto ampio, poiché si estende fino a r uguale a 148, valore al quale scompare. Al suo posto compare un ciclo limite; poi, per r compreso fra 166 e 233, un altro attrattore strano, al quale segue un nuovo ciclo limite. Ciò dimostra la grande varietà di comportamenti dinamici che un modello pur così semplice è in grado di mostrare. Due fisici del Commissariat à l'Énergie Atomique, Yves Po-

Б. П. Белоусов

Описываемая ниже реакция представляет интерес потому, что при проведении ее в водной реакционной смеси возникает ряд скрытых, упорядоченных, окислительно-восстановительных процессов, один из которых через определенный промежуток времени периодически выявляется отчетливым временным изменением цвета всей взятой реакционной смеси. Такое чередующееся изменение окраски от бесцветной до желто-бурой, и наоборот, наблюдается в течение 1 часа и больше, если составные части раствора были взяты в определенных количествах и в соответствующем разведении.

Например, периодическое изменение окраски можно наблюдать при комнатной температуре в 10 *мл* водного раствора следующего состава: лимонная кислота — 2,00 *г*; сульфат церия — 0,16 *г*; бромат калия — 0,20 *г*; серная кислота (1 : 3) — 2,0 *мл*; вода до общего объема — 10,0 *мл*.

Увеличение концентрации водородных ионов или церия, а также повышение температуры весьма ускоряет ритм реакции.

Механизм этой реакции может быть представлен на основании следующих соображений:

В первую очередь должна протекать реакция окисления лимонной кислоты 4-валентным церием

La memoria originale di Belousov, pubblicata nel 1959, descrive la scoperta della reazione chimica più curiosa che oggi si conosca. L'inizio del testo suona così: «La reazione descritta nel seguito è interessante perché, se la si realizza in ambiente acquoso, compare una serie di processi regolari di tipo ossidoriduttivo, uno dei quali si traduce dopo un certo tempo in una variazione periodica del colore della soluzione. Questo cambiamento di colore, dall'incolore al giallo e viceversa, si osserva per un'ora o anche più, se si usano quantità sufficienti di reagenti in concentrazioni opportune».

meau e P. Manneville, ne hanno dato di recente un altro esempio, dimostrando che il modello di Lorenz è anche sede di un fenomeno d'intermittenza quando si passa dal primo ciclo limite al secondo attrattore strano. Quando *r* raggiunge e poi supera il valore critico 166, la soluzione ottenuta è in effetti costituita da zone di oscillazione regolare inframmezzate di quando in quando da ondate di comportamento caotico. Di tempo in tempo il sistema abbandona dunque il ciclo limite per visitare l'attrattore strano e tornare poi al ciclo limite. La transizione verso la turbolenza segue in questo caso una strada diversa da quella osservata nel flusso di Couette o nell'instabilità di Rayleigh-Bénard. Questo risultato, che può suggerire o orientare gli esperimenti, illustra molto bene l'interesse euristico del modello di Lorenz e, più in generale, delle simulazioni numeriche.

Qualunque sia il contributo dei metodi di simulazione, e anche prescindendo dal rischio inerente di ottenere risultati artificiosi, è da ultimo l'esperimento che deve confermare, confutare o decidere. La cinetica delle reazioni chimiche, che come si sa obbedisce a equazioni differenziali in genere non lineari, potrebbe dunque rivelarsi, al pari dell'idrodinamica, un campo d'indagine appropriato. La cosa in verità non è tanto immediata, poiché le reazioni chimiche conosciute per lo più consumano via via certe specie, dette reagenti, per formarne altre chiamate prodotti. Un'evoluzione monotona è dunque la regola generale, ma per fortuna vi sono alcune eccezioni. Le più frequenti riguardano reazioni durante le quali un processo fisico (diffusione di materia o di calore, formazione di una pellicola superficiale e così via) entra in competizione con la trasformazione chimica stessa e le velocità dell'uno e dell'altra evolvono in senso contrario senza mai riuscire a equilibrarsi. Ne risulta un comportamento periodico nel tempo, tutto sommato simile al funzionamento di un vaso di Tantalo. Due esempi noti da tempo sono la luminescenza del fosforo e la dissoluzione periodica dei metalli negli acidi. Altre reazioni, molto più rare, evolvono del pari in maniera non monotona, ma senza l'intervento di alcun processo fisico. In questo caso le oscillazioni sono dovute soltanto a interazioni fra le diverse specie chimiche presenti. Il prototipo per eccellenza di una situazione del genere è la reazione di Belousov-Žabotinskij, che ora descriveremo più in dettaglio. Per il suo carattere particolarissimo, questo genere di reazione è il candidato naturale, se mai ve ne sono, allo studio delle dinamiche complesse, compresa la turbolenza.

La reazione di Belousov-Žabotinskij

Nel 1959 il chimico russo B. P. Belousov scoprì per caso delle oscillazioni nel corso della reazione di ossidazione dell'acido citrico da parte degli ioni BrO_3^- in ambiente acido, catalizzata dalla coppia redox Ce^{3+}/Ce^{4+}. A questa singolare osservazione Belousov consacrò un breve articolo in una rivista assai poco diffusa («Collections of Abstracts on Radiation Medicine», Mosca), sicché la sua scoperta passò inosservata. Per la cronaca si può notare che questa memoria è oggi una delle più citate e, senza dubbio, delle meno lette, perché è difficilissimo procurarsela anche in Unione Sovietica (*si veda l'illustrazione nella pagina a fronte*). Belousov, che morì poco tempo dopo, non andò al di là di questo primo risultato, che doveva tuttavia dargli una celebrità postuma. Lo studio della reazione fu ripreso, o meglio intrapreso, da altri ricercatori e in particolare da A. M. Žabotinskij, che ne fece l'argomento della sua tesi. Più tardi, di questa reazione, singolarissima sotto molti aspetti, si interessarono anche scienziati occidentali, in primo luogo il danese H. Degn. In effetti il comportamento della reazione è per lo meno insolito, sotto il profilo sia temporale sia spaziale. Mantenendo uniforme l'ambiente di reazione tramite un'agitazione continua, si osserva tutta una varietà di evoluzioni differenti purché le condizioni di concentrazione delle specie chimiche, di *p*H e di temperatura siano state scelte in modo opportuno: per esempio si possono avere oscillazioni regolarissime che, se vengono mantenute in un reattore a flusso, possono fungere da meccanismo regolatore di un vero e proprio orologio chimico, esattamente come il bilanciere di un orologio da polso o il pendolo di un orologio a pesi. Si possono ottenere facilmente altri regimi periodici o quasiperiodici più complessi, di cui riparleremo. La singolarità di questa reazione non si limita però a questo e, benché ciò non abbia legami diretti con gli studi condotti attualmente sulla turbolenza chimica, vale la pena di ricordare altre due proprietà che permettono di cogliere tutto l'interesse scientifico di questo sistema di reazione. Effettuata in un reattore a flusso continuo sotto agitazione, la reazione di Belousov-Žabotinskij può, in un certo campo di valori del flusso di alimentazione, dar luogo a un fenomeno di multistabilità: ciò significa che il sistema può trovarsi in due stati diversi per un'identica scelta delle condizioni sperimentali. Lo stato adottato dipende dalla storia precedente dell'ambiente di reazione, che è quindi in grado di costituire una memoria. Questa situazione implica l'esistenza di una vera e propria isteresi chimica. La seconda proprietà, ancora più strana e di cui la reazione di Belousov-Žabotinskij offre l'unico esempio finora conosciuto in chimica, è la seguente. Partendo da una miscela uniforme di reagenti, dopo qualche minuto compare (naturalmente in assenza di agitazione) una strutturazione dello spazio di reazione sotto forma di anelli, se la soluzione è contenuta in una provetta, e di cerchi concentrici o di spirali, se forma uno strato sottile. A livello fondamentale ciò significa che nel corso dell'evoluzione la simmetria dell'ambiente di reazione si abbassa da sola e che dunque la disomogeneità nasce spontaneamente dall'omogeneità. Si deve a P. Glansdorff e a Ilya Prigogine (dell'Université Libre de Bruxelles) la dimostrazione che in questo fenomeno non c'è nulla che contraddica il secondo principio della termodinamica, non appena il sistema si trovi abbastanza lontano dal suo stato di equilibrio. Nonostante tutto, è affa-

La turbolenza chimica, come ogni altro regime dinamico della reazione di Belousov-Žabotinskij, può essere studiata sotto il profilo quantitativo con l'ausilio di questo dispositivo. Una pompa peristaltica alimenta con reagenti prelevati da serbatoi un reattore aperto, termostatato e posto sotto agitazione. Un fascio di luce monocromatica (λ = 340 nanometri) permette di raccogliere all'uscita del fotomoltiplicatore un segnale di ampiezza direttamente proporzionale alla concentrazione di una delle specie chimiche presenti (Ce^{4+}). Dopo l'amplificazione il segnale misurato è introdotto in forma numerica, a intervalli di tempo regolari, nella memoria di un calcolatore che controlla l'esperimento. Il risultato di parecchie migliaia, o decine di migliaia, di misurazioni viene così registrato in un archivio di dati che può essere poi sottoposto a opportune elaborazioni numeriche per individuare le principali caratteristiche dinamiche della reazione chimica.

scinante osservare come un sistema chimico così semplice possa essere sede di un'evoluzione tanto ricca e svariata.

Attraverso questi fenomeni di autorganizzazione del tempo e dello spazio, la reazione di Belousov-Žabotinskij si presenta come un nuovo intermediario tra il mondo della fisica-chimica e quello dei sistemi biologici. Ciò basta a spiegare i numerosi lavori che da alcuni anni le vengono dedicati. Benché non si conosca con certezza il suo meccanismo in tutti i particolari, se ne possiede tuttavia un modello semplificato con tre specie chimiche intermedie che sembra spiegare abbastanza bene le sue proprietà essenziali. Questo modello comporta una fase autocatalitica, cioè una reazione nel corso della quale una specie chimica aumenta la propria velocità di produzione proporzionalmente alla quantità già presente: è una sorta di «effetto valanga» chimico. Ebbene, si è dimostrato che questa caratteristica, o più in generale l'intervento di un processo di retroazione, è una condizione necessaria, ma non sufficiente, perché compaia una struttura. Dato che la retroazione è un meccanismo molto diffuso nei sistemi biologici e invece poco frequente nei sistemi chimici, è naturale che i fenomeni di autorganizzazione si osservino nei primi più che nei secondi. A questo riguardo la reazione di Belousov-Žabotinskij ha un evidente valore paradigmatico, tanto più che, nonostante tutti gli sforzi compiuti, finora non è stato scoperto alcun altro sistema chimico radicalmente diverso da essa e capace di imitarla. Questa reazione è diventata dunque il caso per eccellenza su cui si concentrano numerose ricerche interdisciplinari; lo studio della turbolenza chimica aggiunge semplicemente una corda al suo arco.

La turbolenza chimica

L'idea di ricorrere alla dinamica chimica per studiare la turbolenza debole, avanzata come abbiamo ricordato da Ruelle nel 1973, si è concretizzata nel corso degli ultimi 12 o 13 anni quando, al di là dei problemi contingenti, gli sperimentatori sono riusciti a superare un ostacolo concettuale che non sarebbe giusto sottovalutare a posteriori. È stato cioè necessario abbandonare il proposito di interessarsi alla reazione di Belousov-Žabotinskij in sé e considerarla semplicemente un nuovo supporto per le ricerche sulla turbolenza, cosa che per i chimici non era affatto naturale, tanto più che, per adottare questo punto di vista, bisognava concentrare l'attenzione su osservazioni che fino a quel momento erano state sistematicamente scartate a causa del loro carattere caotico e non riproducibile, attribuito di solito a carenze sperimentali. Ammesso questo, il resto venne quasi da sé, e i lavori sulla fluidodinamica servirono in un primo momento da guida. I nomi di Degn e di Otto E. Rössler in Europa, di R. Schmitz e di J. L. Hudson negli Stati Uniti sono, fra gli altri, legati a studi qualitativi preliminari, mentre il gruppo da noi diretto al Centre de Recherche Paul Pascal di Bordeaux ha effettuato un'analisi spettrale quantitativa analoga a quella descritta prima a proposito del flusso di Couette e della convezione di Rayleigh-Bénard. A questo scopo, la reazione di Belousov-Žabotinskij viene effettuata sotto agitazione in un reattore termostatato e alimentato con un flusso continuo di portata regolabile (*si veda l'illustrazione a pagina 149*). La densità ottica $D(t)$ della soluzione a 340 nanometri permette di caratterizzare la dinamica di reazione, perché una sola specie chimica assorbe la luce di questa lunghezza d'onda, lo ione Ce^{4+}. Questa proprietà ha qui lo stesso ruolo che la velocità istantanea ha nelle esperienze di fuidodinamica.

Quando la portata del flusso di alimentazione aumenta l'evoluzione della trasformata di Fourier di $D(t)$ è molto interessante, purché le condizioni di temperatura e di concentrazione siano scelte opportunamente. Quando la portata è relativamente scarsa, la reazione oscilla in modo molto regolare: lo spettro di Fourier contiene una sola frequenza fondamentale f_1 perfettamente individuata (*si veda* a *nell'illustrazione di pagina 144*). Poi, quando il flusso di alimentazione aumenta, compare un secondo modo di oscillazione di frequenza f_2; vi è una biforcazione da un regime periodico verso un regime quasiperiodico (*si veda* b *nella medesima illustrazione*). In seguito si osserva un fenomeno di agganciamento tra le frequenze f_1 e f_2 che porta a un regime quasiperiodico degenere (*c*). Infine nello spettro di Fourier emerge una banda larga (*f*) prima che il sistema assuma, per valori elevati del flusso di alimentazione, uno stato stazionario.

Quest'ultimo risultato non ha nulla di sorprendente, perché le specie chimiche non hanno in pratica più il tempo di reagire tra loro, dato che la loro permanenza nel reattore è troppo breve. Si potrebbe quindi dire che l'insorgere della turbolenza chimica segue passo passo le stesse tappe della turbolenza fluidodinamica. Questi esperimenti non forniscono però solo una conferma in più alla teoria di Ruelle e Takens: in effetti, poiché il segnale $D(t)$ è pochissimo deteriorato dal rumore (è il frutto infatti di una misura semplicissima, quasi rozza), l'analisi numerica può essere spinta oltre. In particolare, l'esame della funzione di autocorrelazione del segnale (funzione che caratterizza la rapidità con la quale il sistema «perde il ricordo» dei suoi stati precedenti) completa utilmente le informazioni fornite dalla trasformata di Fourier. Grazie a ciò è possibile stabilire senza ambiguità se la comparsa della frequenza f_2 corrisponda a una modulazione della frequenza o della fase del modo f_1 e non a una modulazione della sua ampiezza. Alcuni diranno certamente che si tratta di un dettaglio tecnico, tuttavia ciò testimonia i progressi realizzabili sul piano sperimentale.

Un altro esempio, molto più appariscente, è fornito dallo studio dell'attrattore che regola la dinamica.

Un attrattore strano

Via via che elementi sempre più numerosi confermano la teoria di Ruelle e Takens, la ricerca del tipo di attrattore e la sua rappresentazione si rivelano un obiettivo del tutto naturale. I sistemi modello, come quello di Lorenz, non presentano in questo senso alcuna difficoltà particolare, dato che tutte le variabili dinamiche, e di conseguenza lo spazio delle fasi, sono definite dal modello stesso. L'integrazione numerica fornisce direttamente i valori di queste variabili e basta lasciare che una traiettoria si sviluppi per un tempo abbastanza lungo per osservare l'attrattore che via via si costituisce. Naturalmente quando si passa all'ambito sperimentale le cose vanno in modo diverso. In questo caso, lungi dal poter conoscere o soltanto enumerare l'insieme delle variabili dinamiche pertinenti, ci si accontenta per ora, non potendo fare di più, di misurarne una. Così per individuare il meccanismo della reazione di Belousov-Žabotinskij vengono prese in considerazione una ventina o più di specie chimiche intermedie; e tuttavia non vi è la sicurezza che ciò sia sufficiente, o tanto meno esauriente. Quante sono fra queste le specie indipendenti, ossia i gradi di libertà? Nessuno lo sa. Quindi la determinazione delle variazioni di concentrazione dello ione Ce^{4+} fornisce solo un'informazione molto parziale e la rappresentazione dell'attrattore nello spazio delle fasi ci sfugge senza rimedio. In fluidodinamica accade la stessa cosa.

Malgrado tutto si deve sottolineare che sono le proprietà topologiche dell'attrattore, come per esempio la sensibilità alle condizioni iniziali, a costituire il vero polo di interesse, sicché un'immagine dell'attrattore che conservi queste proprietà è più che sufficiente. Il problema che si pone è quindi quello di trovare un metodo che consenta di elaborare un'immagine siffatta a partire dalla sola registrazione temporale posseduta, $X(t)$ [ovvero $V(t)$ o $D(t)$ negli esempi descritti sopra]. A questo scopo i matematici hanno proposto di ricorrere a uno spazio a tre dimensioni in cui le coordinate di un punto dovrebbero essere: $X(t)$, $X(t + t_1)$, $X(t + t_2)$, dove t_1 e t_2 sono due istanti arbitrari; oppure X, $\dot{X}$, $\ddot{X}$, dove $\dot{X}$ e $\ddot{X}$ rappresentano le derivate prima e seconda del segnale $X(t)$ rispetto al tempo. Questa proposta equivale ad assimilare, sotto il profilo topologico, $X(t + t_1)$, $X(t + t_2)$, $\dot{X}$ o $\ddot{X}$ a variabili dinamiche indipendenti da X. Alcuni calcoli compiuti su un sistema modello dimostrano che ciò non è affatto irragionevole. Pur mantenendo una riserva sulla validità di questa proposta, si può allora pensare di ricostruire la topologia dell'attrattore che regola la dinamica in uno spazio semplificato, diverso dallo spazio delle fasi. Perché quest'operazione riesca, si deve tuttavia supporre che in pratica la proprietà misu-

rata sia nota con una precisione relativa molto grande, altrimenti le incertezze della misurazione danno luogo a un groviglio inestricabile di traiettorie da cui non si può ricavare nulla. Sotto questo profilo la reazione chimica si presenta in una luce molto favorevole. La prova è fornita dall'illustrazione di pagina 142, che rappresenta l'attrattore, o meglio la sua immagine in uno spazio $X, \dot{X}, \ddot{X}$. L'attrattore, come si vede, dà luogo a un'orbita ben definita (*in alto a destra nell'illustrazione*) che si tuffa in un vortice. Quando una traiettoria penetra in questo vortice, vi compie un numero di giri variabile prima di uscirne, cosicché è impossibile sapere in anticipo per quanto tempo il punto rappresentativo resterà all'interno. Ciò esprime chiaramente la sensibilità alle condizioni iniziali. Consideriamo in effetti due punti vicini a un istante dato: a seconda che la loro traiettoria successiva entri o non entri nel vortice e a seconda del numero di giri che vi compie, la distanza che li separa varierà in modo considerevole e in maniera non prevedibile. Per quanto ne sappiamo, è la prima volta che un'osservazione di questo genere è compiuta direttamente sulla base di risultati sperimentali.

A questa prima illustrazione sperimentale del concetto di attrattore strano si affiancano prospettive a breve termine che giustificano le speranze basate sulla dinamica chimica. In primo luogo si pensa di riuscire a mettere in luce altre caratteristiche proprie degli attrattori strani e a introdurre una sorta di valutazione quantitativa della turbolenza. I teorici hanno già fatto in questo senso diverse proposte, che sono state più o meno collaudate su modelli; ma la loro applicazione sperimentale non ha ancora avuto esito. Restano poi da scoprire altri comportamenti tipici. Abbiamo potuto osservare nella reazione di Belousov-Žabotinskij il fenomeno di intermittenza di cui si è già parlato a proposito del modello di Lorenz. Altre ricerche sono in corso per trovare e analizzare una demoltiplicazione di frequenza in cascata in condizioni sperimentali più accessibili di quelle in cui sono stati effettuati i lavori sull'elio liquido. Jack Turner dell'Università del Texas ad Austin è riuscito a dimostrare tramite simulazione numerica che questo fenomeno deve con tutta probabilità presentarsi. Il suo calcolo è basato su uno schema semplificato del meccanismo di reazione, che fa intervenire solo tre specie intermedie. Proposto nel 1974 da Richard M. Noyes e Richard J. Field dell'Università di Eugene nell'Oregon, da cui il nome di «Oregonator» con cui è divenuto celebre, questo modello spiega piuttosto bene numerose osservazioni relative alla reazione di Belousov-Žabotinskij e ciò gli conferisce un grado di credibilità molto notevole. Infine i sistemi chimici potrebbero fornire precisazioni sull'evoluzione del comportamento dinamico nelle vicinanze immediate dei punti di biforcazione, informazioni decisive per giudicare la validità di certe previsioni teoriche.

Il bilancio dei lavori sulla turbolenza debole dimostra che in questo settore sono stati compiuti di recente vari progressi, e ne vengono compiuti ancora quasi ogni giorno. Sono emerse alcune idee guida, in particolare la distinzione tra regime quasiperiodico e regime non periodico, idee che hanno implicazioni molto profonde. I metodi teorici e sperimentali concordano sull'esistenza di almeno tre itinerari che portano alla comparsa della turbolenza debole: il passaggio per un regime quasiperiodico, la demoltiplicazione in cascata di una frequenza oppure un fenomeno d'intermittenza. Tutto ciò non deve peraltro farci dimenticare che resta molta strada da compiere per comprendere la turbolenza e la sua genesi. Certo la risposta ai problemi posti dall'insorgere della turbolenza debole non sembra più fuori dalla nostra portata, e si può nutrire una ragionevole speranza di acquisire una visione abbastanza generale del comportamento caotico associato ai sistemi dinamici non lineari. Ma per il momento nulla dimostra che i sistemi con pochi e quelli con molti gradi di libertà evolvano in modo simile e che sia possibile studiarli con gli stessi metodi di analisi. Ciò indica che la soluzione di questo grande problema della fisica richiede ancora molti sforzi.

Il mescolamento dei fluidi

Esperimenti e modelli al calcolatore consentono di spiegare come fluidi viscosi che scorrono in due dimensioni con moto periodico possano generare il caos necessario a un efficiente mescolamento

di Julio M. Ottino

Che cosa hanno in comune l'eruzione del Krakatoa, la preparazione della pasta sfoglia e la luminosità delle stelle? Tutti e tre hanno in comune una qualche forma di mescolamento. Un violento mescolamento di magmi potrebbe avere innescato l'eruzione del Krakatoa; spianando e ripiegando l'impasto (un tipico processo di mescolamento) si ottiene la sfoglia, mentre è il mescolamento che avviene all'interno di una stella a determinarne la composizione chimica e quindi la luminosità superficiale. Esempi di mescolamento si possono trovare letteralmente in tutto l'universo in un grandissimo intervallo di scale di tempi e di dimensioni. I gas emessi si mescolano con l'aria dell'ambiente in pochi secondi, mentre i processi di mescolamento che hanno luogo nel mantello terrestre possono durare anche parecchie centinaia di milioni di anni.

Il mescolamento ha un ruolo critico anche nella moderna tecnologia. Gli ingegneri chimici si basano sul mescolamento per assicurarsi che le sostanze reagiscano adeguatamente per produrre miscele di polimeri dotate di proprietà particolari e per disperdere nelle condutture agenti che riducano la resistenza. Tuttavia, nonostante la sua ubiquità in natura e nell'industria, il mescolamento è conosciuto in modo incompleto. In effetti, i ricercatori non possono neppure basarsi su una terminologia comune: il mescolamento è spesso chiamato «rimescolamento» in oceanografia e in geofisica, «miscelazione» in ingegneria chimica e «agitazione» nell'ingegneria dei processi produttivi.

Indipendentemente dal nome, non vi è dubbio che il mescolamento sia un processo di estrema complessità e sia presente in una grande varietà di sistemi. Per esempio, per costruire una teoria del mescolamento dei fluidi si devono prendere in considerazione fluidi che possono essere miscibili o parzialmente miscibili, reagenti o inerti, e flussi che sono lenti e ordinati oppure molto veloci e turbolenti. Non sorprende perciò che una teoria unica non sia in grado di spiegare tutti gli aspetti del mescolamento dei fluidi e che semplici calcoli di routine non riescano solitamente a catturare tutti i particolari importanti.

Eppure, sia esperimenti fisici, sia simulazioni al calcolatore possono fornire informazioni sul processo di mescolamento. Negli ultimi anni i miei colleghi e io abbiamo seguito entrambi i metodi nel tentativo di migliorare la comprensione di vari aspetti del processo, in particolare del mescolamento relativo a flussi lenti e a fluidi viscosi quali gli oli.

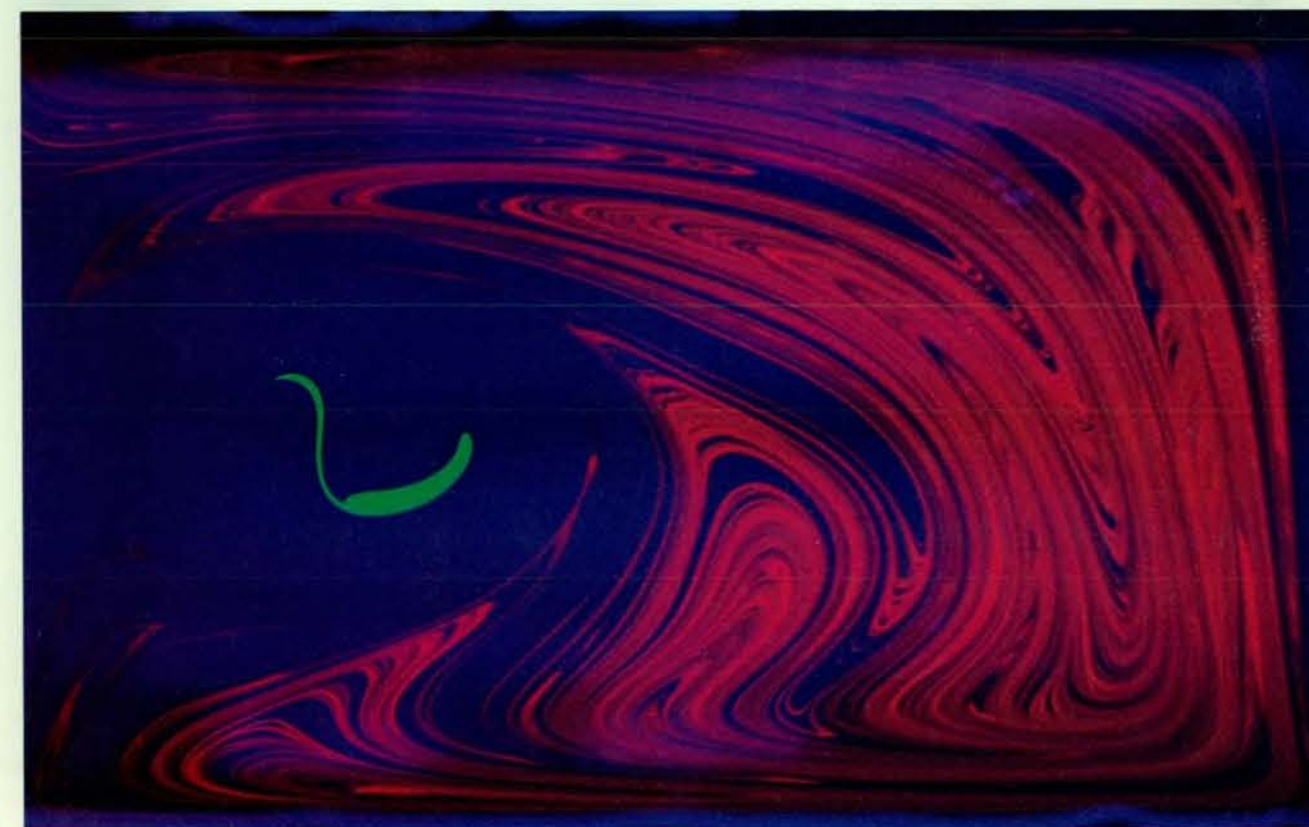
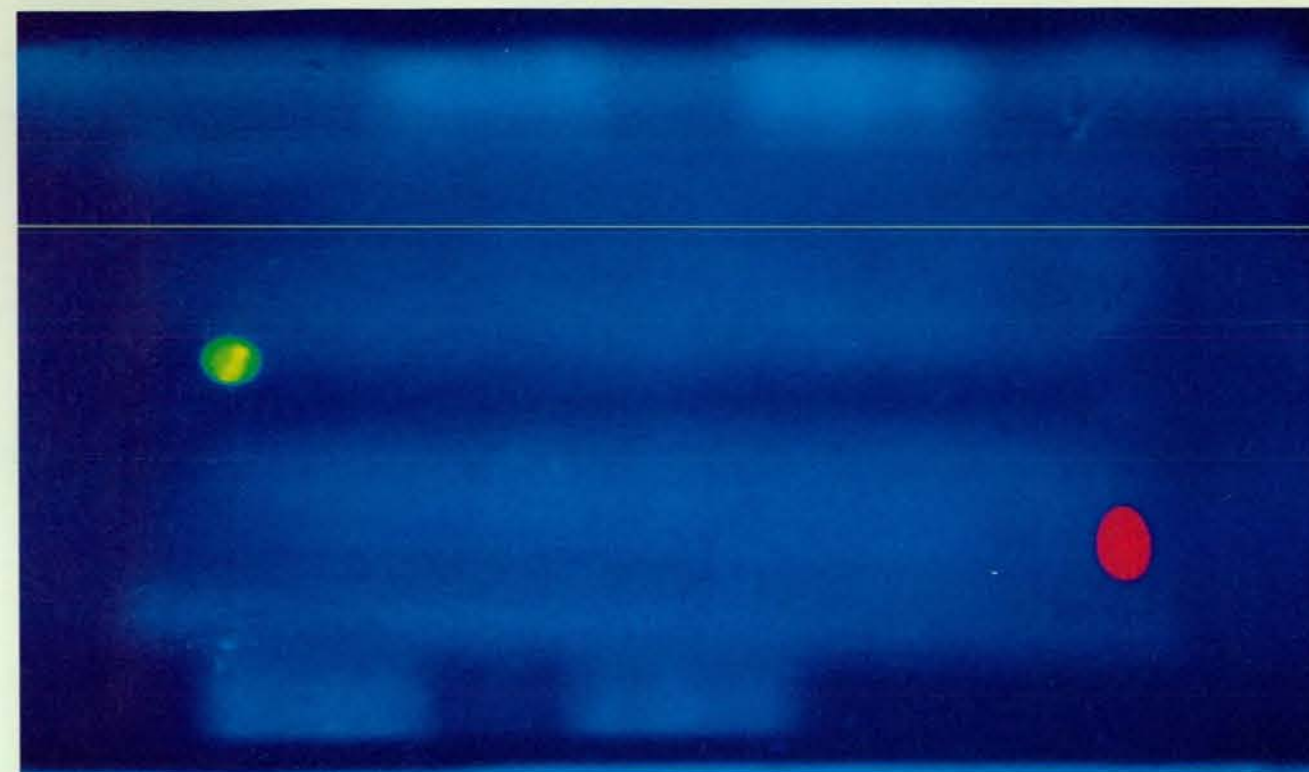
Unire insieme due colori a olio è un buon esempio di mescolamento di fluidi viscosi. Dopo solo pochi secondi di mescolamento si può produrre una sconcertante figura di striature allungate e ripiegate. (I rilegatori sfruttano il fenomeno nella «marmorizzazione» che orna talvolta le copertine o i risguardi dei libri.) Tuttavia, a meno che non si sia voluto intenzionalmente il contrario, è possibile trovare alcune «isole» di colore non mescolato fra striature a forma di spirale. Anche se il mescolamento di fluidi viscosi può produrre strutture fantasticamente complesse, esso può anche produrre figure con un certo grado di regolarità e coerenza.

I miei studenti e io all'Università del Massachusetts ad Amherst abbiamo cercato di caratterizzare i flussi che producono tali figure eseguendo esperimenti e simulazioni al calcolatore che si rifanno al mescolamento di due colori. In alcuni dei nostri esperimenti iniettiamo «gocce» di glicerina colorata in una massa di glicerina incolore contenuta in una cavità profonda. Quando le pareti della cavità vengono fatte muovere periodicamente, le forze di taglio che esse esercitano sul fluido viscoso contenuto nella cavità possono allungare e ripiegare la goccia colorata in modo piuttosto complesso; l'intera cavità mostra presto una complessa figura di pieghe su pieghe. Contemporaneamente però può accadere che nello stesso contenitore una goccia analoga non subisca alcun allungamento; può darsi che la goccia si muova e ruoti, ma poi torni regolarmente alla posizione iniziale. Perché si producono figure tanto diverse?

Fondamenti di meccanica dei fluidi

La chiave per comprendere gli aspetti fondamentali di questo tipo di mescolamento sta nel concetto di «moto», un'idea che si può far risalire al matematico svizzero del XVIII secolo Leonhard Euler. Il moto di un fluido è un'espressione matematica che stabilisce dove si troverà in qualsiasi istante futuro ogni particella del fluido. Se potessimo conoscere il moto di un determinato flusso, si potrebbe in linea di principio conoscere quasi tutto ciò che c'è da sapere sul mescolamento che esso produce. Per esempio, si potrebbero calcolare le forze e l'energia totale necessarie per ottenere un certo grado di mescolamento nel sistema.

Durante il secolo scorso la descrizione del flusso in termini di moto di un fluido è stata in gran parte abbandonata a favore di una descrizione basata sul campo delle velocità del fluido: un'espressione che specifica la velocità del fluido in ogni punto della regione di flusso in un istante qualsiasi. Però, se si conosce il moto, è facile calcolare il campo delle velocità, mentre la conoscenza del campo delle velocità non è sufficiente per calcolare esplicitamente il moto. Essendo quindi il moto in un certo senso una descrizione maggiormente basilare del flusso, i miei collaboratori e io abbiamo preferito lavorare con quello che forse molti considerano un concetto superato.

Strettamente connesso al concetto di moto è quello di trasformazione puntuale, un'operazione matematica che consente di identificare una particella di fluido e di specificarne la posizione in ogni istante futuro. Ogni particella di fluido viene «mappata» in una nuova posizione

Sia i flussi caotici sia quelli non caotici sono evidenti in un esperimento eseguito da Kenny Leong e dall'autore nel loro laboratorio all'Università del Massachusetts ad Amherst. Una cavità rettangolare viene riempita di glicerina e due «gocce» di tracciante fluorescente (una verde e una rossa) sono iniettate al di sotto della superficie (*in alto*). Ogni parete della cavità può scorrere parallelamente a se stessa indipendentemente dalle altre. In questo esperimento le pareti superiore e inferiore vengono fatte muovere periodicamente, ma in modo discontinuo. La parete superiore si sposta da sinistra verso destra per un certo tempo, quindi si ferma; a questo punto la parete inferiore si sposta alla stessa velocità e per lo stesso intervallo di tempo, ma da destra verso sinistra. I due spostamenti costituiscono un periodo. Dopo 10 periodi (*in basso*) la goccia rossa, situata in una regione di mescolamento caotico, è stata allungata e ripiegata varie volte. La goccia verde è stata allungata solo in misura ridotta: essa evidenzia un'«isola» di mescolamento non caotico.

applicando la trasformazione. Particelle identificate inizialmente come separate non possono occupare la stessa posizione nello stesso istante e una particella non può dividersi in due. Anche se una trasformazione puntuale esiste in teoria per tutti i flussi di mescolamento, essa si può ottenere esattamente solo per i casi più semplici. Per questo motivo, la maggior parte di ciò che si sa sul mescolamento è limitata a flussi relativamente semplici, come i flussi lineari nei quali le linee di un tracciante non si piegano. Eppure è probabile che questi tipi di flusso non riescano a far individuare i processi che portano a un efficiente mescolamento, i quali sono per loro natura non lineari. Per avere almeno un'idea di ciò che entra in gioco in tali processi, si devono considerare flussi stazionari in due dimensioni.

Flussi bidimensionali

Tutti i flussi bidimensionali sono costituiti dagli stessi elementi: punti iperbolici (chiamati anche di sella) e punti ellittici (*si veda l'illustrazione a pagina 39*). Nel caso di punto iperbolico il fluido si muove verso di esso lungo una direzione e se ne allontana lungo un'altra. Invece nel caso di punto ellittico il fluido circola attorno al punto. (Dovrei anche ricordare che esiste un terzo tipo di punto, il punto parabolico, nel quale il moto del fluido è di taglio, ossia tangenziale. Tali punti si trovano, per esempio, in un fluido che scorre lungo una parete solida. Si possono trascurare i punti parabolici nella descrizione della natura del mescolamento nei flussi bidimensionali.) Come è prevedibile, il mescolamento in un flusso stazionario bidimensionale è piuttosto inefficiente in confronto al mescolamento in flussi tridimensionali, in particolare in quelli che cambiano continuamente nel tempo. Infatti, due sono le possibilità in un flusso stazionario bidimensionale in una regione limitata: o le particelle di fluido seguono ripetutamente le stesse traiettorie, le cosiddette linee di flusso, oppure non si muovono affatto.

Dal momento che nei flussi stazionari le linee di flusso sono fisse e le traiettorie delle particelle di fluido non possono mai incrociarsi, le particelle di fluido non hanno alcuna possibilità di venire reciprocamente in contatto, ossia di mescolarsi. Esiste un modo per liberarsi del confinamento delle linee di flusso cosicché le particelle di fluido possano evitare di dover seguire ripetutamente la stessa linea di flusso? La risposta è positiva, se si può modificare nel tempo il moto del fluido in modo che una linea di flusso a un istante intersechi quella relativa a un istante successivo.

Il modo più semplice per farlo (e il più facile per analizzarlo) è di costringere il flusso a variare periodicamente nel tempo. Perché un tale flusso porti a un mescolamento efficiente, però, esso deve essere in grado di estendere e ripiegare una regione di fluido e di riportarla - allungata e ripiegata - alla sua posizione iniziale. L'operazione di estensione e ripiegamento corrisponde a quella che viene chiamata mappa a ferro di cavallo, descritta da Stephen Smale dell'Università della California a Berkeley.

Il fatto che per poter mescolare efficientemente una sostanza si debba riportarne una parte nella sua posizione iniziale è contro ogni intuizione. Eppure, se il mescolamento avviene in un sistema limitato non vi è realmente alcuna alternativa. Se si lancia ripetutamente una freccia su un bersaglio, qualche lancio finirà con il cadere arbitrariamente vicino a qualche altro, dal momento che la superficie del bersaglio è semplicemente finita. Per lo stesso motivo, la ripetizione di allungamenti e di ripiegamenti in una cavità chiusa porterà immancabilmente, in certi istanti, particelle di fluido vicinissime alle loro posizioni iniziali.

Se in un flusso periodico una particella di fluido ritorna dopo un certo tempo nella sua esatta posizione iniziale, la particella definisce quello che viene detto un punto periodico. A seconda del numero di periodi necessari perché la particella ritorni alla propria posizione di partenza, si parla di punto periodico di periodo uno, di periodo due, e così via. Un punto periodico si può anche classificare come iperbolico o ellittico a seconda della direzione del flusso nelle sue immediate vicinanze.

Quando un punto periodico ellittico descrive la sua traiettoria ciclica, il materiale circostante non solo circola attorno a esso (come farebbe attorno a un punto ellittico fisso), ma si muove anche con esso. Tuttavia, nonostante la rotazione e la traslazione del materiale, questo non cede facilmente materia nella restante parte del flusso. Tali regioni di materiale sono viste come «isole» di fluido e il mescolamento all'interno delle isole è tipicamente lento. Dato che il materiale non può né entrare né abbandonare le vicinanze di un punto periodico ellittico, tali punti costituiscono ostacoli a un efficiente mescolamento.

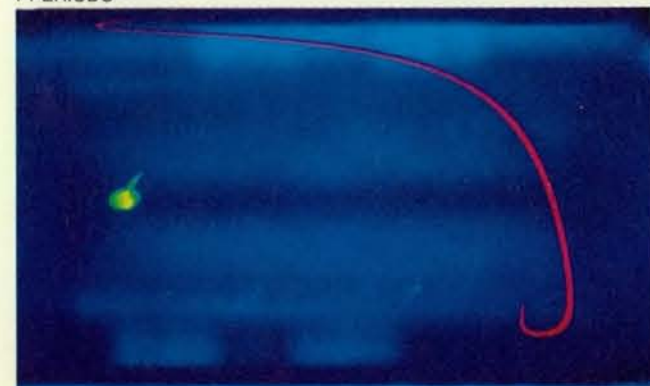
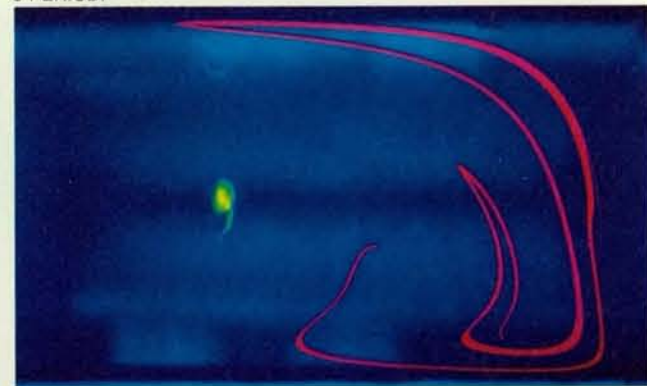
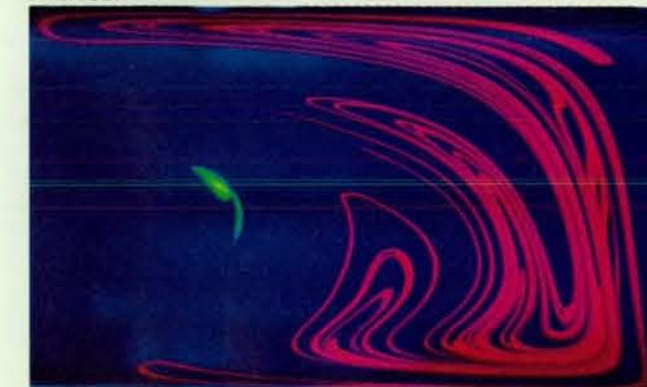
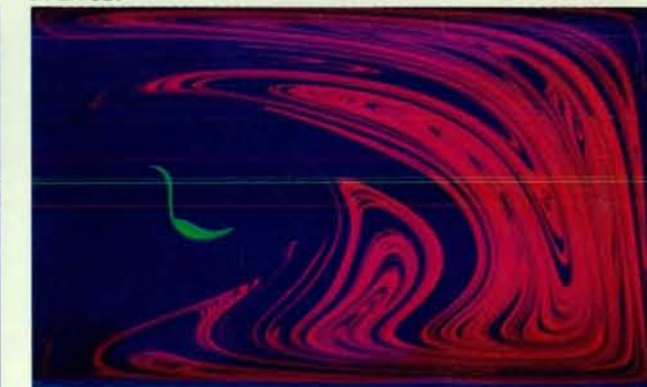
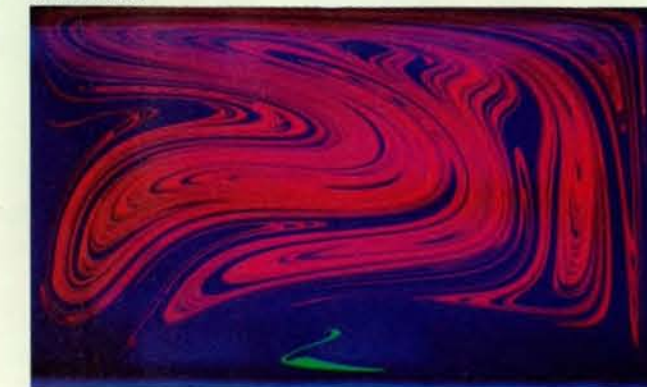
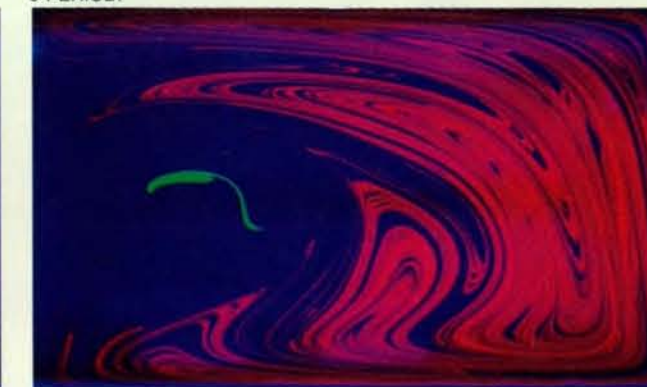
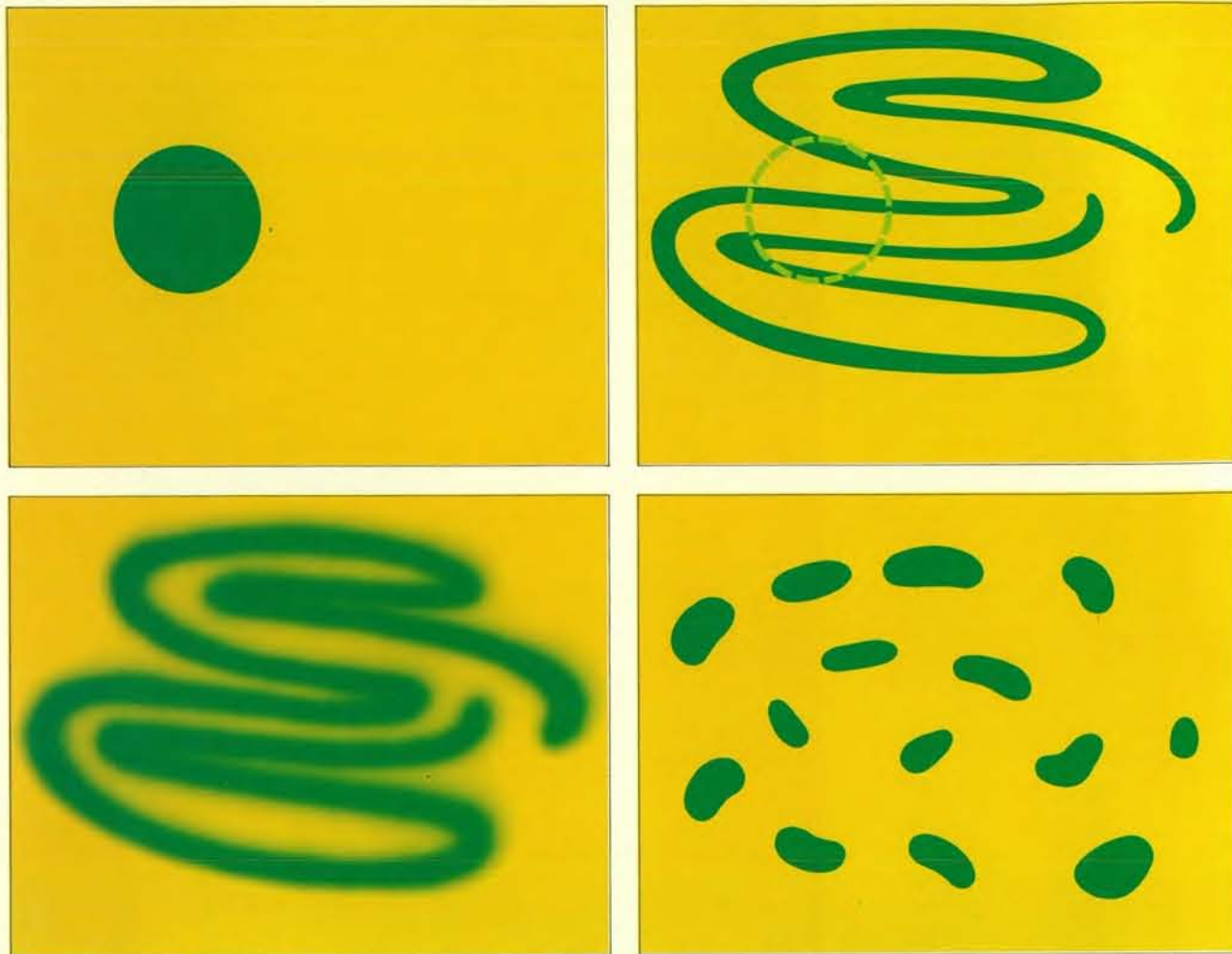
Del pari, mentre un punto periodico iperbolico descrive la sua traiettoria ciclica, il materiale circostante che si muove con il punto subisce una contrazione in una direzione e un allungamento in un'altra. Nel fare ciò il punto spande filamenti allungati di fluido in una direzione e attira materiale in un'altra. (Se supponiamo che i fluidi siano incomprimibili, l'allungamento e la contrazione devono compensarsi a vicenda.)

Le impronte digitali del caos

Dove va a finire il materiale sparso da un punto periodico iperbolico? Da dove viene il materiale che si avvicina al punto? Può darsi che un afflusso si fonda in modo regolare con un efflusso, cioè che il materiale sparso da un punto iperbolico venga attirato dallo stesso o da un altro punto iperbolico. Questo è proprio ciò che accade in un flusso stazionario (sebbene in tal caso i punti iperbolici siano fissi e non periodici), con il risultato che il flusso non deforma e non ripiega il materiale in maniera efficiente.

I flussi bidimensionali dipendenti dal tempo possono però tradursi in efficienti meccanismi di allungamento e di ripiegamento poiché in tali flussi è possibile che una regione di efflusso associata a un punto periodico iperbolico attraversi la regione di afflusso dello stesso o di un altro punto iperbolico. Un punto nel quale l'afflusso e l'efflusso di un solo punto iperbolico si intersecano viene detto «punto omoclino trasversale». Se l'intersezione deriva da flussi di due punti iperbolici diversi, esso viene detto «punto eteroclino trasversale».

Le intersezioni omocline ed eterocline sono le impronte digitali del caos. Da un punto di vista matematico, allora, un sistema in grado di produrre mappe a ferro di cavallo o intersezioni trasversali omocline o eterocline si può classificare come caotico. Ne consegue che una mappa a ferro di cavallo implica in realtà l'esistenza di punti trasversali omoclini; analogamente l'identificazione di uno solo di tali punti è sufficiente a implicare l'esistenza di una mappa a ferro di cavallo.

Il fatto che un solo attraversamento di afflusso e di efflusso conduca invariabilmente a punti trasversali omoclini, e che un siffatto attraversamento possa aver luogo perfino in quelli che paiono sistemi fisici «ben educati» descritti dalle leggi del moto di Newton fu scoperto per la prima volta dal matematico francese del XIX secolo Henri Poincaré. Tuttavia l'analisi del comportamento straordinariamente complesso derivante da tali attraversamenti (oggi chiamato «caos») mise

L'allungamento e il ripiegamento tipici del mescolamento caotico sono evidenziati dalla goccia rossa in questa sequenza di fotografie dell'esperimento mostrato nell'illustrazione della pagina precedente. Dopo tre periodi è chiaramente visibile lo schema fondamentale di allungamento e ripiegamento. L'isola verde che contraddistingue una regione di mescolamento essenzialmente non caotico e le pieghe che caratterizzano una regione di mescolamento caotico si muovono nella cavità, ma dopo ogni periodo tornano nelle stesse posizioni (anche se alquanto deformate). I filamenti tracciati dalla goccia verde indicano che essa subisce una rotazione completa ogni due periodi circa. Se si eseguisse l'esperimento all'indietro, la goccia verde ritornerebbe quasi alla forma e alla posizione iniziali, dato che l'errore nella riproduzione dei suoi movimenti in senso inverso cresce linearmente. È invece praticamente impossibile lo «smescolamento» della goccia rossa: in tal caso l'errore nel riprodurre i suoi movimenti in senso inverso cresce esponenzialmente.

Il mescolamento di fluidi nei processi naturali o industriali coinvolge, oltre all'allungamento e al ripiegamento, la diffusione molecolare e la rottura. Un goccia di vernice colorata in una massa fluida (*in alto a sinistra*) può essere allungata e ripiegata indefinitamente senza diffondere o rompersi (*in alto a destra*) solo in un caso ideale. In tale situazione, per ottenere un mescolamento efficiente, una parte della goccia deve tornare alla posizione iniziale. La diffusione molecolare (senza cui è impossibile il totale mescolamento) rende normalmente indistinguibili i confini tra i fluidi miscibili (*in basso a sinistra*). Nel caso di fluidi immiscibili la goccia, nell'allungarsi, può spezzarsi in goccioline che possono successivamente aderire per formare molte gocce più piccole di quella iniziale (*in basso a destra*).

in crisi Poincaré, il quale decise di non esplorare ulteriormente l'argomento.

Nella stessa misura in cui il mescolamento può essere rappresentato da una trasformazione puntuale deterministica, esso dovrebbe essere cinematicamente reversibile. In altre parole, dovrebbe essere possibile «riseparare» i fluidi (perlomeno se si trascura la diffusione molecolare). Tuttavia, l'esperienza quotidiana suggerisce che il mescolamento è un processo irreversibile. Anche se il sistema è in teoria deterministico, i moti che conducono a ripetuti allungamenti e ripiegamenti non si possono invertire.

Una situazione abbastanza simile esiste in altri sistemi fisici, quali quelli analizzati da Poincaré, formati da molte particelle i cui moti sono descritti da equazioni deterministiche. (Questi tipi di sistemi vengono comunemente detti «sistemi hamiltoniani».) Uno dei più famosi fisici americani del XIX secolo, J. Willard Gibbs, scoprì che persino i sistemi hamiltoniani possono avere una irreversibilità e una imprevedibilità intrinseche, e il fatto che egli sia ricorso per spiegarlo a un esperimento ideale relativo al mescolamento dà la misura della sua intuizione. Le sue considerazioni passarono apparentemente inosservate fino a quando l'oceanografo svedese Pierre Wellander non le sottolineò in un acuto articolo pubblicato nel 1955.

Come catturare il caos in un flusso

Il fatto che l'allungamento e il ripiegamento abbiano un ruolo di primo piano nel mescolamento era noto in ingegneria chimica fin dagli anni cinquanta in seguito alla ricerca pionieristica di Robert S. Spencer e Ralph M. Wiley della Dow Chemical Company e di William D. Mohr e collaboratori della E. I. du Pont de Nemours & Company, Inc., mentre le conseguenze dell'esistenza di mappe a ferro di cavallo e di punti omoclini ed eteroclini sono state riconosciute solo recentemente.

Pare che sia stato il matematico russo Vladimir I. Arnold ad aver trovato la prima correlazione diretta tra caos e flussi di fluidi. Secondo Michel Hénon, un astronomo francese dell'Osservatorio di Nizza, Arnold suggerì nel 1965 la possibilità che sistemi fluido-meccanici possano mostrare traiettorie caotiche di particelle. Hénon approfondì l'ipotesi di Arnold e in un articolo di tre pagine che conteneva una sola illustrazione fu capace di dimostrare che un flusso stazionario tridimensionale di un fluido non viscoso può dare effettivamente origine a linee di flusso caotiche.

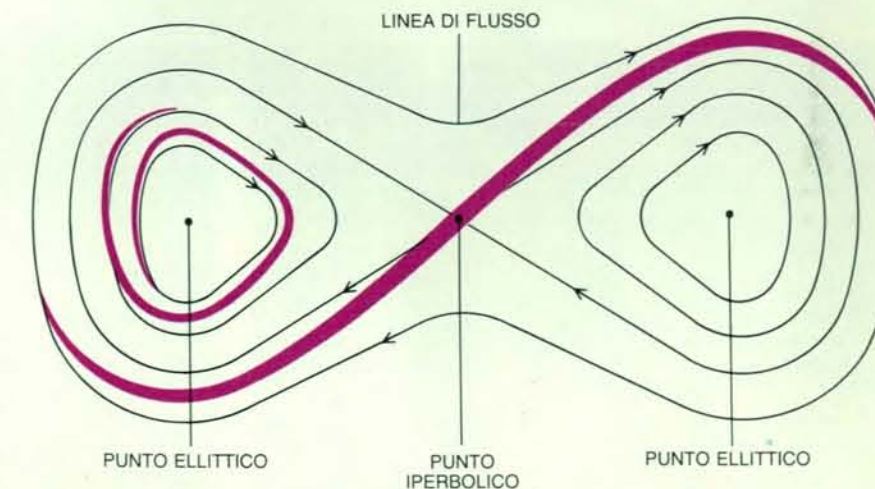
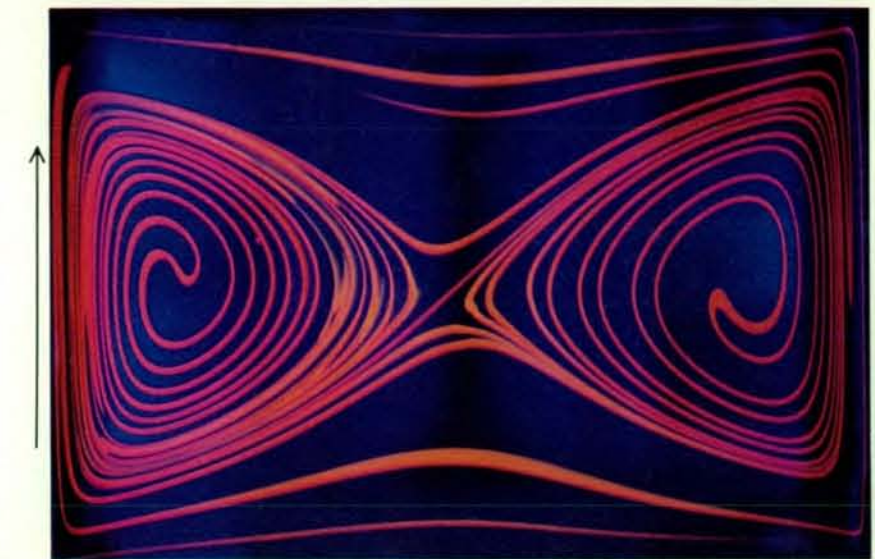
Nel 1984 Hassan Aref, allora alla Brown University, osservò che le equazioni che descrivono le traiettorie delle particelle di fluido in un flusso bidimensionale sono formalmente identiche a quelle che descrivono un sistema hamiltoniano. Egli spinse oltre le sue osservazioni, dimostrando, mediante una simulazione al calcolatore, che un sistema hamiltoniano soggetto a forze periodiche può in realtà produrre un mescolamento efficiente.

A tre dimensioni l'analogia tra mescolamento e sistemi hamiltoniani non funziona, mentre a due dimensioni l'analogia è esatta: il mescolamento di fluidi si può considerare come una rappresentazione visiva del comportamento di un sistema hamiltoniano caotico. La ricerca di Aref, tenendo conto del fatto che in laboratorio è molto più facile studiare flussi bidimensionali che non flussi tridimensionali, mi ha suggerito di cercare tracce di caos in un sistema sperimentale di flusso in una cavità che insieme ai miei studenti ho costruito ad Amherst nel 1983.

Immagini di flussi

Kenny Leong, un mio studente diplomato, e io riuscimmo a determinare le posizioni approssimative di alcuni punti periodici e di strutture di grande scala in flussi bidimensionali registrando accuratamente immagini stroboscopiche del sistema in movimento. (Essendo noi interessati al mescolamento rapido, ci siamo concentrati sul comportamento di punti periodici di ordine basso, cioè con periodo uno, due o tre; i punti di ordine superiore non partecipano al processo con la stessa frequenza di quelli di ordine più basso.) In un tipico esperimento introduciamo gocce di vernice fluorescente in certe posizioni della cavità rettangolare, illuminiamo la cavità con luce ultravioletta, mettiamo in movimento i suoi lati secondo un particolare schema di moto e registriamo le posizioni delle gocce e le relative contorsioni riprendendo fotografie del sistema a intervalli regolari. Se il mescolamento è efficiente, le particelle di vernice esplorano una grande regione del sistema. Se invece il mescolamento è modesto, le gocce cedono il colorante solo lentamente alla massa del fluido oppure rimangono vicine a punti periodici ellittici.

In altri esperimenti Paul D. Swanson, un altro dei miei studenti, e io ci siamo occupati specificamente di flussi che trovano una soluzione analitica esatta delle equazioni di moto del fluido. In tal modo siamo in grado di confrontare al meglio i nostri risultati sperimentali con quelli previsti dalla teoria. Sfortunatamente, il numero di sistemi per i quali esistono soluzioni analitiche esatte è piuttosto piccolo, e molti sono talmente idealizzati da non poter essere riprodotti in esperimenti di laboratorio. Uno dei sistemi che ammette soluzione esatta ed è realizzabile sperimentalmente è il flusso tra due cilindri eccentrici rotanti. Un siffatto sistema è stato studiato anche da Aref (oggi all'Università della California a San Diego) e da Michael Tabor e Rene Chevray della Columbia University.

Esperimenti a tappeto su flussi caotici bidimensionali rivelano che le strutture fluide di mescolamento di grande scala (quali la posizione e la forma di isole e di grandi pieghe) sono perfettamente riproducibili; i più piccoli particolari delle strutture allungate e ripiegate non lo sono. Il motivo dipende dal fatto che piccole deviazioni dalla posizione iniziale delle gocce di vernice vengono amplificate all'interno delle regioni caotiche di flusso. Ed è proprio così che deve essere: dovrebbe essere impossibile ripro-

I punti ellittici e iperbolici sono aspetti tipici di flussi lenti bidimensionali. La fotografia (*in alto*), eseguita da Leong e dall'autore, mostra uno di tali flussi, prodotto quando le pareti opposte di una cavità rettangolare riempita di glicerina vengono spostate in senso opposto con velocità costante. Le linee in arancione (prodotte da un tracciante iniettato all'inizio lungo una linea che va dall'angolo a sinistra in basso a quello a destra in alto) sono quasi allineate con le linee di corrente del flusso, cioè le linee descritte nel flusso stazionario dalle particelle di fluido in moto. L'andamento di flusso contiene tre punti fissi: un punto centrale iperbolico e due punti ellittici ai lati. Il flusso nelle vicinanze di ogni punto ellittico (*in basso*) produce un vortice che ruota in senso orario; esso fa aumentare linearmente con il tempo la lunghezza del tracciante. Il flusso nelle vicinanze di un punto iperbolico si avvicina al punto in una direzione e se ne allontana in un'altra. Poiché il materiale del fluido non può attraversare le linee di flusso, un flusso stazionario bidimensionale non è efficiente nel mescolamento. Tuttavia, se si fa variare il flusso nel tempo, i filamenti allungati del tracciante non hanno sufficiente tempo per allinearsi con nuove linee di flusso e vengono rapidamente ripiegati da un cambiamento di direzione del flusso.

durre esattamente qualsiasi fase dei nostri esperimenti di mescolamento. In fondo, obiettivo del mescolamento è creare una distribuzione casuale. Ed è esattamente ciò che si ottiene con i meccanismi di allungamento e di ripiegamento impiegati negli esperimenti.

È anche interessante notare come il caos possa pacificamente coesistere con simmetrie nel flusso, quali quelle rappresentate da punti periodici. Infatti, eliminando sistematicamente le simmetrie in un flusso caotico, i miei collaboratori e io siamo riusciti ad aumentare l'efficienza di mescolamento del flusso.

Sperimentazione ed elaborazione a confronto

Se il sistema sperimentale è piuttosto semplice (tale da potersi ricavare un'espressione matematica del campo delle velocità), esso può essere facilmente simulato con un calcolatore. In un programma tipico viene strategicamente disposto un certo numero di «particelle» di prova in un moto simulato, ossia in un campo delle velocità; le posizioni calcolate delle particelle dopo circa 1000 periodi offrono in tal caso un buon quadro del comportamento generale del sistema dopo che è stato messo in funzione per molto tempo. L'immagine ottenuta con questo tipo di simulazione è detta «sezione di Poincaré» e una sezione di Poincaré di aspetto complesso viene spesso assunta come prova al calcolatore dell'esistenza del caos (*si veda la parte superiore dell'illustrazione della pagina a fronte*). Le simulazioni del mescolamento al calcolatore presentano anche una forma di irreversibilità cinematica, ma nel loro caso essa deriva dall'ingrandimento esponenziale degli errori introdotti dal calcolatore, che può operare soltanto con numeri formati da un numero finito di cifre.

Se esistono simulazioni di mescolamento al calcolatore, perché preoccuparsi di eseguire esperimenti fisici? La prima cosa da tenere presente è che la risoluzione del campo delle velocità deve essere molto più elevata per simulazioni di mescolamento che per la maggior parte degli altri problemi di meccanica dei fluidi.

Anche campi delle velocità piuttosto semplici possono dare origine a strutture estremamente complesse (*si vedano le illustrazioni alle pagine 36 e 37*); in alcuni problemi di mescolamento sarebbe interessante risolvere alcuni dei particolari più fini delle strutture.

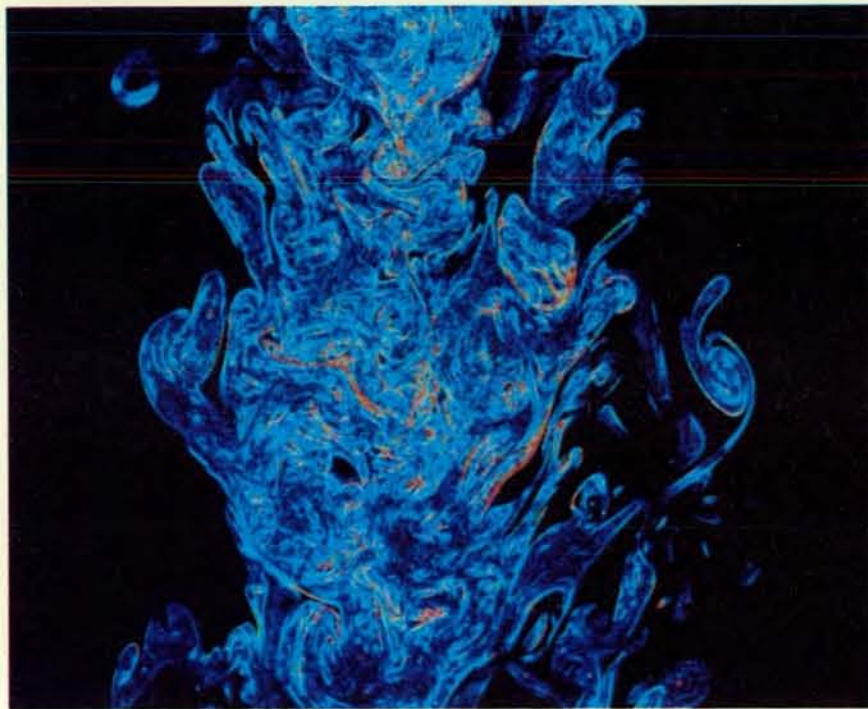
Per esempio, in una simulazione del flusso in una cavità rettangolare un campo delle velocità calcolato convenzionalmente potrebbe rivelarsi troppo grezzo per poter evidenziare i particolari delle striature allungate e ripiegate. Sarebbe anche virtualmente inutile per individuare le esatte posizioni dei punti periodici che determinano il comportamento complesso dei flussi caotici. Inoltre, mentre nella maggior parte dei problemi di meccanica dei fluidi l'obiettivo è quello di ottenere un'approssimazione del campo delle velocità, nel mescolamento il problema inizia, anziché terminare, con la specificazione del campo delle velocità.

Per tale motivo gli studi sul mescolamento sono stati in gran parte focalizzati su quelle che sono in realtà caricature di flussi (descritte da equazioni che in alcuni casi si possono risolvere esattamente) anziché su problemi più realistici (le cui soluzioni possono essere soltanto approssimate). In effetti, i metodi numerici per approssimare le soluzioni delle equazioni della meccanica dei fluidi introducono spesso effetti spuri che non esistono veramente nei problemi reali di mescolamento dei fluidi.

Comunque simulando con il calcolatore i flussi semplificati dei nostri esperimenti si incontrano spesso difficoltà insormontabili. Il calcolatore tratta un fluido come se fosse formato da elementi discreti. In una simulazione centinaia di migliaia di tali elementi possono costituire una singola goccia di vernice e il numero di calcoli necessari per tracciare una mappa del suo comportamento caotico in un flusso di mescolamento può essere enorme.

Per registrare tutte le striature in regioni di mescolamento caotico in un esempio relativamente semplice, quale quello mostrato nelle illustrazioni delle pagine 36 e 37, sarebbero necessari 300 anni di tempo macchina di un calcolatore capace di effettuare un milione di ope-

Il modello a frullatore, ideato da John G. Franjione e dall'autore, serve a illustrare il processo fondamentale di allungamento e ripiegamento caratteristico del mescolamento (*a*). Una linea tracciata sulla superficie di una cella di fluido viene allungata e ripiegata quando una lama del frullatore avanza nel fluido prima in direzione perpendicolare alla linea (*b*) e poi in direzione parallela a essa (*c*). La linea viene allungata senza spezzarsi; qualsiasi parte che si estenda oltre la sommità della cella rientra dal basso, mentre le parti che si estendono oltre la parete di sinistra rientrano dalla parete di destra. Un calcolatore può produrre immagini della cella che dipendono dal numero di volte in cui le lame hanno attraversato la cella. Nelle immagini mostrate una sola linea iniziale formata da 100 000 punti è stata allungata e ripiegata 16 volte in condizioni di mescolamento differenti. Il mescolamento che ne risulta può essere confinato a regioni della cella (*d*) o può estendersi a tutta la cella (*e*), a seconda del «vigore» con il quale le lame agitano il fluido.

Il flusso viscoso in un cuscinetto, ovvero il flusso tra due cilindri rotanti eccentrici, può essere simulato al calcolatore. Se i cilindri vengono fatti muovere periodicamente in versi opposti, il flusso conduce a un mescolamento caotico, come si vede nella sezione di Poincaré del sistema per 1000 periodi (*in alto*) e nella sua mappa di allungamento per 10 periodi (*in basso*). Una sezione di Poincaré viene prodotta inserendo un certo numero di «particelle» colorate di prova nel flusso che simula un sistema di mescolamento, calcolando il loro moto per ogni periodo e spostandole nelle loro nuove posizioni. Una mappa di allungamento indica le regioni nelle quali il fluido è stato allungato in un flusso simulato. La maggior parte dell'allungamento avviene all'interno delle regioni bianche, mentre nelle regioni colorate l'allungamento è modesto. La mappa di allungamento mostrata è sorprendentemente simile alla struttura prodotta nel flusso reale (*si veda l'illustrazione di copertina*). Le immagini sono state realizzate ad Amherst da Paul D. Swanson e dall'autore.

Il flusso turbolento può generare strutture molto diverse da quelle prodotte in un flusso viscoso lento. L'immagine, ottenuta da K. R. Sreenivasan della Yale University, è una ricostruzione al calcolatore di un getto d'acqua espulso da un ugello circolare in acqua quieta. Le strutture del flusso venivano originariamente registrate su pellicola sciogliendo un colorante fluorescente nell'acqua espulsa e dirigendo una lama di luce laser lungo l'asse dell'ugello. L'intensità della fluorescenza risultante è proporzionale al gradiente di concentrazione relativo del colorante nell'acqua; le immagini sono state codificate in colore dal blu scuro al rosso a seconda del gradiente di concentrazione. Il flusso turbolento mostrato appare formato da varie strutture frattali sovrapposte, tra cui numerosi vortici.

razioni in virgola mobile al secondo. A dire il vero, si potrebbe ritenere che non sia necessaria un'indagine così dettagliata e che sarebbe meglio una valutazione statistica dell'allungamento. Ma non sarebbe un'ammissione di incapacità? Se si può conoscere esattamente il campo delle velocità (o il moto), perché si dovrebbe affrontare il problema dal punto di vista statistico?

Per concludere, è necessario associare sviluppi teorici a esperimenti ben progettati, poiché - con molta probabilità - le tecniche di elaborazione da sole non sono in grado di rispondere a molte domande sui flussi caotici.

Per esempio, a quali tipi di moto devono essere soggette le pareti di una cavità per poter ridurre le dimensioni di tutte le isole in essa contenute (comprese quelle nuove che potrebbero comparire spontaneamente) al di sotto di un determinato livello? La risposta a questa domanda potrebbe un giorno consentire la progettazione di un sofisticato sistema di riconoscimento delle forme che sia in grado di rilevare la presenza di isole in un sistema di mescolamento e di alterare successivamente il flusso in modo tale da mescolare le isole al resto del fluido.

Limitazioni e complicazioni

Prima però di poter costruire questo tipo di macchine «intelligenti» di mescolamento, si dovrà saperne molto di più sui flussi reali. Anche se gli esperimenti e le simulazioni al calcolatore descritti in questo articolo forniscono qualche informazione su problemi generali del mescolamento (per esempio, come aumentare esponenzialmente la superficie di contatto tra due fluidi), essi rappresentano casi di problemi ideali piuttosto particolari. Per fare un esempio, i flussi in cavità descritti in questo articolo non presentano inerzia. In altre parole, il flusso cessa non appena le pareti della cavità smettono di muoversi. Di conseguenza tali flussi non portano ad alcuno dei processi caratteristici osservati nei flussi turbolenti.

Per dirlo in termini più tecnici, i numeri di Reynolds (il rapporto fra le forze inerziali e quelle viscose in un fluido) dei flussi studiati nei nostri esperimenti erano bassi. Flussi caratterizzati da bassi numeri di Reynolds (i cosiddetti flussi laminari) sono ordinati e regolari, mentre quelli caratterizzati da numeri di Reynolds elevati producono campi delle velocità variabili nel tempo piuttosto complessi che portano a un rapido mescolamento. Un osservatore in un punto fisso della nostra cavità sperimentale vedrebbe ripetersi periodicamente lo stesso semplice campo delle velocità anziché vedere i campi non periodici e imprevedibili prodotti in un flusso turbolento. Tuttavia è proprio a causa della turbolenza che è più facile mescolare la panna nel caffè con un cucchiaino (un sistema con un numero di Reynolds relativamente alto) che mescolare due vernici colorate per interni con una spatola (un sistema con un numero di Reynolds basso).

Pur avendo in un certo senso escluso dalla trattazione i flussi di mescolamento più efficaci (quelli turbolenti), vi è motivo di credere che alcune idee presentate in questo articolo possano nondimeno portare a concetti utili per lo studio di tali flussi. Per esempio, versioni lievemente più elaborate di flussi caotici bidimensionali mostrano una velocità non periodica quando sono misurate in un punto determinato. È però chiaramente necessaria una ricerca molto più approfondita prima che si possa comprendere la turbolenza altrettanto bene di quanto oggi conosciamo i flussi laminari.

In questa trattazione ho anche semplificato l'argomento supponendo che nel mescolamento la diffusione non sia importante. Ma le cose non stanno realmente così. Per prendere in considerazione l'effetto della diffusione nel mescolamento si può ricorrere a un semplice modello che ipotizza che la velocità di diffusione tra due striature contigue di due sostanze miscibili venga controllata dalla rapidità con la quale le striature vengono «strizzate» e assottigliate, la quale dipende a sua volta dalla componente del flusso in una direzione perpendicolare alle striature. In questo modo il mescolamento ha un doppio effetto che accelera la diffusione: fa aumentare la superficie di contatto tra i fluidi, riducendo nel contempo la distanza attraverso la quale i fluidi devono diffondere e aumentando i gradienti di concentrazione. Un tale modello si può in pratica estendere fino a includere l'effetto del mescolamento su reazioni chimiche come la combustione.

Un altro processo comune che - per semplicità - ho ignorato è la rottura delle goccioline nei fluidi immiscibili, che è in realtà un fenomeno molto complesso. Vi sono due casi limite: un fluido ad alta viscosità disperso in una massa di fluido a bassa viscosità e un fluido a bassa viscosità disperso in una massa di fluido ad alta viscosità. Entrambi i casi sono difficili da analizzare, ma per ragioni differenti. Nel primo caso il fluido a bassa viscosità è sottoposto all'insieme degli sforzi di taglio dal momento che esso non può trasmettere efficacemente le sollecitazioni alle goccioline del fluido ad alta viscosità. In realtà un flusso stazionario con sforzi di taglio non può spezzare una gocciolina che ha una viscosità circa quattro volte maggiore di quella del fluido che la tiene in sospensione. Sotto questo aspetto hanno più successo i flussi con grandi allungamenti di quelli con sforzi di taglio. Però i primi non potrebbero essere particolarmente efficaci nel caso in cui le goccioline a bassa viscosità siano disperse in un fluido ad alta viscosità, poiché è necessario allungare considerevolmente le goccioline prima di romperle.

I miei collaboratori e io abbiamo eseguito studi sperimentali riguardanti il mescolamento di due fluidi di diversa viscosità. Come era prevedibile, l'entità delle rotture è molto minore all'interno delle isole che nelle regioni di caos. Del resto, potrebbe capitare che un eccessivo rimescolamento possa fare aderire le goccioline fra loro; i fluidi potrebbero talvolta separarsi a causa della loro coalescenza. Usando semplici modelli al calcolatore, siamo stati in grado di prevedere la cinetica di tale aggregazione in flussi caotici semplici. Infine - e con una chiara evidenza - resta il fatto che tutti gli esperimenti da noi finora eseguiti riguardavano flussi bidimensionali, laddove nel mondo reale ci si trova a eseguire misurazioni in tre dimensioni. Solo da poco i miei studenti e io abbiamo costruito la prima apparecchiatura capace di produrre esperimenti di mescolamento controllato in flussi tridimensionali e stiamo cominciando ora i relativi esperimenti. Vi sono molte domande di base riguardanti il mescolamento in flussi tridimensionali lenti e sfortunatamente alcune delle intuizioni che abbiamo ricavato dal nostro studio dei flussi bidimensionali non si estendono necessariamente ai flussi tridimensionali.

Le striature caratteristiche del mescolamento che avviene in un flusso viscoso sono evidenti in questa roccia magmatica della catena vulcanica Inyo della California orientale. In effetti, la roccia è il prodotto del mescolamento di due magmi differenti, uno dei quali (quello caratterizzato dalle striature più chiare) conteneva microscopiche bollicine di sostanze volatili. La diffusione attraverso tali striature del magma è molto lenta; il tempo necessario perché la diffusione cancelli striature di spessore dell'ordine di un centimetro è superiore all'età della Terra. La fotografia è di Ichiro Sugioka e Bradford Sturtevant del Caltech.

Il primo passo di un lungo viaggio

L'elenco dei problemi di mescolamento non finisce qui. Il mescolamento di fluidi viscoelastici (fluidi, quale il Silly Putty, che ritornano alla loro forma originaria dopo essere stati deformati) è un problema formidabile sul quale si sa poco, nonostante il fatto che esso emerga in modo prominente nella lavorazione di polimeri ad alto peso molecolare. Il mescolamento di fluidi delicati, non in grado di sopportare variazioni rapide degli sforzi di taglio senza essere degradati, è importante in bioingegneria. Il mescolamento di fluidi altamente viscosi dovuto a moti termici interessa i geofisici che studiano il mescolamento dei magmi nel mantello terrestre.

Nonostante la scoraggiante complessità dei processi di mescolamento, si spera che possano essere compresi e che tali conoscenze possano essere poi applicate proficuamente in laboratorio e negli impianti chimici. Inoltre, poiché semplici esperimenti fungono da analoghi del caos, essi potrebbero chiarire alcuni aspetti fondamentali di sistemi caotici in generale.

BIBLIOGRAFIA

AREF H., *Stirring by Chaotic Advection* in «Journal of Fluid Mechanics», 143, giugno 1984.

RANZ WILLIAM E., *Fluid Mechanical Mixing-Lamellar Description* in «Mixing of Liquids by Mechanical Agitation», a cura di Jaromir J. Ulbrecht e Gary K. Patterson, Gordon and Breach Science Publishers, 1985.

KHAKHAR D. V., RISING H. e OTTINO J. M., *Analysis of Chaotic Mixing in Two Model Systems* in «Journal of Fluid Mechanics», 172, pp. 419-451, novembre 1986.

OTTINO J. M., LEONG C. W., RISING H. e SWANSON P. D., *Morphological Structures Produced by Mixing in Chaotic Flows* in «Nature», 333, n. 6172, 2 giugno 1988.

Il caos nel mondo microscopico

I vetri di spin e lo studio dei mezzi disordinati

di J. Hammann e M. Ocio

I vetri sono dei solidi ma, a differenza che nei cristalli, gli atomi che li costituiscono non sono disposti in strutture che si ripetono identiche riempiendo lo spazio: in un vetro le posizioni degli atomi possono formare localmente strutture più o meno regolari, che però non si ripetono indefinitamente. Per analogia, la locuzione «vetri di spin» designa sistemi magnetici in cui le orientazioni dei momenti magnetici elementari (gli spin degli atomi) non si ripetono su grandi distanze. Mentre in un corpo ferromagnetico gli spin sono paralleli e in un corpo antiferromagnetico due spin vicini sono antiparalleli, in un vetro di spin le orientazioni degli spin sono aleatorie.

L'orientazione dei momenti magnetici degli ioni dipende dalla loro interazione. A bassa temperatura, per effetto di queste interazioni, gli spin di un qualunque sistema si orientano gli uni rispetto agli altri in una disposizione che rende minima la loro energia d'interazione. Così nella struttura ferromagnetica i momenti magnetici sono tutti paralleli e orientati nello stesso verso. Nella struttura antiferromagnetica gli spin sono paralleli tra loro, ma sono rivolti alternatamente in verso opposto. Nello stato di vetro di spin le orientazioni degli spin sono completamente aleatorie.

Lo studio di questo stato magnetico si colloca in un quadro generale: quello delle indagini sulla materia disordinata, che oggi è il principale polo di attrazione della fisica dello stato solido. Sono molti i fattori che spiegano l'attrattiva esercitata dai sistemi aleatori. Il più evidente è l'onnipresenza di queste strutture: in natura i mezzi disordinati sono di gran lunga i più diffusi, e sono anche i meno conosciuti. Da un punto di vista pratico, i composti amorfi semiconduttori o ferromagnetici sono molto largamente utilizzati e negli ultimi anni l'industria delle materie plastiche e dei polimeri vetrosi ha conosciuto un grande sviluppo.

A un livello più basilare i modelli teorici che descrivono le proprietà dei sistemi ben ordinati, come i cristalli, benché non possano mai essere risolti esattamente a causa del gran numero di unità che costituiscono un solido, possono essere trattati correttamente con metodi statistici, nei quali si sfruttano le simmetrie che caratterizzano un reticolo cristallino. Le approssimazioni utilizzabili in questi casi ordinati sono ben definite e il loro ambito d'applicazione è perfettamente noto. Le cose non stanno così per i sistemi disordinati, dove l'assenza di simmetria rende notevolmente complessa la trattazione teorica. Gli sviluppi recenti della meccanica statistica e i moderni strumenti informatici di calcolo hanno aperto tuttavia vaste possibilità: benché l'intuizione suggerisca che le proprietà della materia disordinata debbano essere affatto imprevedibili, la teoria, e in particolare quella che è stata sviluppata nello studio dei vetri di spin, ha dimostrato che in realtà si può definire un comportamento medio e che esiste in un certo senso un ordine nascosto grazie al quale è possibile spiegare un gran numero di proprietà comuni a sistemi di tipi diversi.

Il magnetismo ha sempre fornito alla fisica dello stato solido modelli particolarmente adatti agli studi teorici, soprattutto perché il sistema magnetico è semplice e «ben definito», nel senso che i suoi parametri pertinenti sono noti. Un esempio chiarissimo è dato dall'elaborazione della teoria delle transizioni di fase, che si è valsa in larga misura dello studio sperimentale dei composti ferromagnetici o antiferromagnetici. Lo stesso vale per la fisica dei mezzi aleatori, che nei vetri di spin ha trovato modelli semplici in grado di illustrare le caratteristiche essenziali di questi mezzi. In tutta la descrizione teorica dei vetri di spin è sufficiente conoscere le interazioni aleatorie tra momenti magnetici localizzati. È proprio la semplicità di questa descrizione a renderla generale: in questa ottica possono essere affrontati numerosi casi meno facili da studiare in laboratorio. In particolare, alla fine dell'articolo, ricorderemo la sua applicazione alle reti di neuroni, il cui modello a «vetro di spin» ne illustra la capacità di memorizzare e di riconoscere le forme. Sotto il profilo sperimentale, infine, le proprietà magnetiche sono piuttosto facili da misurare grazie alle moderne tecniche di magnetometria, anche se nei vetri di spin i segnali da rilevare sono molto più deboli che non nei casi classici di ferromagnetismo.

Abbiamo finora utilizzato le locuzioni «sistemi magnetici» e «parametri pertinenti» senza precisare i concetti corrispondenti, cosa che ora faremo. Un composto magnetico è formato da un insieme di atomi, alcuni dei quali possiedono un momento magnetico; il sistema magnetico associato a questo composto è l'insieme di questi momenti, dotati di proprietà specifiche determinate dalla matrice in cui si trovano. I parametri pertinenti sono i parametri necessari alla descrizione delle proprietà magnetiche del sistema. Ciascuno spin si comporta come un dipolo magnetico che tende ad allinearsi con il campo magnetico circostante per effetto di una coppia proporzionale all'intensità di questo campo e al valore del momento del dipolo. Il campo agente sul momento magnetico di un atomo è la somma del campo esterno applicato e di un campo «effettivo» creato dai momenti magnetici di tutti gli altri atomi.

A temperatura non nulla l'agitazione termica provoca una continua riorientazione degli spin e, se la temperatura è elevata, l'effetto delle interazioni fra gli spin è trascurabile rispetto all'agitazione termica: in ciascun punto, il valore medio rispetto al tempo dello spin (vettore che in seguito chiameremo polarizzazione) è nullo. Il sistema è allora in una fase paramagnetica: l'applicazione di un campo esterno, per esempio quello generato da un magnete, crea in ogni punto una polarizzazione parallela alla direzione del campo magnetico esterno. L'anisotropia introdotta è proporzionale all'intensità del campo applicato e inversamente proporzionale alla temperatura.

Se la temperatura è abbastanza bassa,

l'effetto delle interazioni prevale sull'agitazione termica e gli spin assumono una polarizzazione la cui orientazione dipende dalla natura di queste interazioni: il sistema appare sotto una nuova fase (ferromagnetica o antiferromagnetica). Esaminiamo ora questo meccanismo nel contesto dei vetri di spin.

I materiali che in origine hanno ricevuto la denominazione di vetri di spin sono leghe a base di un metallo nobile, come l'oro, l'argento o il rame, e contenenti una piccola percentuale di un metallo magnetico (come ferro o manganese): gli esempi più citati sono quelli delle leghe rame-manganese o argento-manganese. I momenti magnetici degli atomi di manganese sono distribuiti in modo aleatorio nella matrice del metallo nobile e l'interazione fra di essi avviene soprattutto tramite gli elettroni di conduzione. In funzione della distanza che li separa, l'interazione fra due momenti *m* e *m'* può essere ferromagnetica, e allora *m* e *m'* tendono ad assumere una orientazione parallela, oppure antiferromagnetica, e allora *m* e *m'* tendono ad assumere un'orientazione antiparallela. La posizione aleatoria dei momenti magnetici provoca di conseguenza una distribuzione aleatoria delle loro interazioni.

Le caratteristiche dei vetri di spin

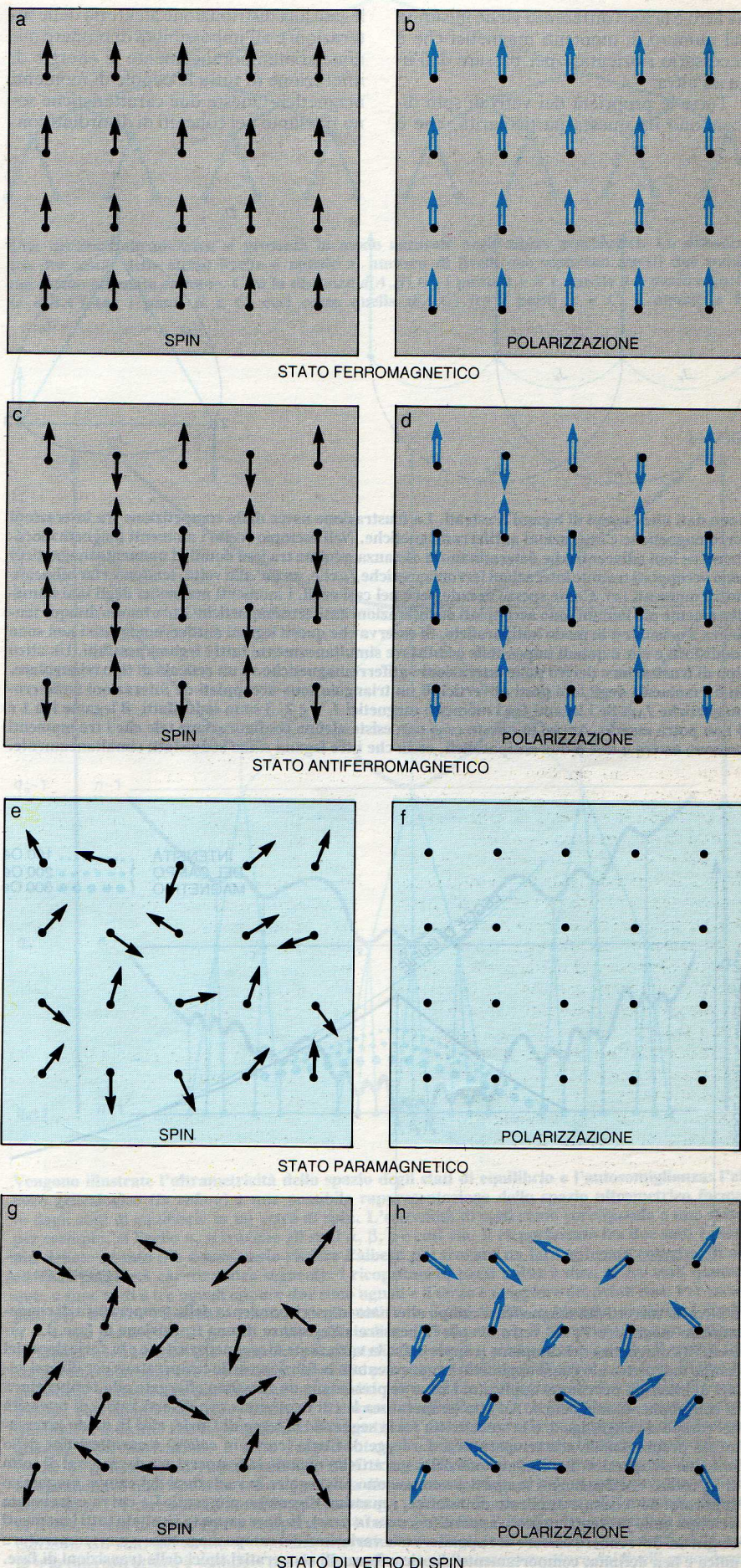
In queste condizioni non è possibile minimizzare simultaneamente le energie di interazione di tutte le coppie di momenti magnetici. Una configurazione è definita da una determinata orientazione di ciascun momento magnetico; a bassa temperatura, quando le fluttuazioni termiche sono deboli, il sistema tende ad adottare la configurazione corrispondente all'energia minima possibile, ma non esiste alcuna configurazione in cui le energie di legame di tutte le coppie di momenti magnetici siano simultaneamente minime; sono numerosissime quindi le configurazioni corrispondenti a energie molto vicine tra loro e vicine all'energia minima. Queste configurazioni possono essere molto diverse

STATO FERROMAGNETICO

STATO ANTIFERROMAGNETICO

STATO PARAMAGNETICO

STATO DI VETRO DI SPIN

Sono illustrate le differenze nella distribuzione dei momenti magnetici e nella polarizzazione relativamente agli stati ferromagnetico, antiferromagnetico, paramagnetico e di vetro di spin. La colonna di sinistra si riferisce alle distribuzioni istantanee dei momenti magnetici (spin); quella di destra alle polarizzazioni, cioè le medie temporali dell'orientazione degli spin. Nello stato ferromagnetico gli spin (*a*), rappresentati a bassa temperatura, e le polarizzazioni (*b*) sono tutti paralleli. Nello stato antiferromagnetico a bassa temperatura spin (*c*) e polarizzazioni (*d*) sono paralleli, ma orientati a due a due in verso opposto. Nello stato paramagnetico l'orientazione degli spin è aleatoria (*e*), ma il loro moto fa sì che le polarizzazioni elementari (*f*) siano nulle. Nello stato di vetro di spin (*g*) i momenti magnetici hanno, a bassa temperatura, un'orientazione fissa perché il magnetismo del sistema è «congelato» e le polarizzazioni (*h*) sono non nulle.

tra loro e la loro differenza viene misurata dal numero di momenti magnetici che è necessario riorientare per passare dall'una all'altra.

Tutte le proprietà dei vetri di spin discendono da questa particolarità, che è legata alla distribuzione aleatoria delle interazioni e all'impossibilità di rendere minime contemporaneamente le energie di interazione di tutte le coppie di momenti magnetici. Queste due caratteristiche sono riassunte nei concetti di disordine congelato e di frustrazione, che hanno consentito di generalizzare il problema e di moltiplicare gli esempi di composti che sono vetri di spin: un composto può presentare le proprietà di un vetro di spin se disordine e frustrazione coesistono. Il disordine dev'essere «congelato», cioè l'inserimento aleatorio dei momenti magnetici nello spazio non deve modificarsi nel tempo, quali che siano la temperatura e l'intensità del campo magnetico. La frustrazione può essere ottenuta in vari modi: abbiamo visto quella che si manifesta nelle leghe intermetalliche; nei composti isolanti, che non possiedono elettroni di conduzione, le interazioni tra momenti magnetici sono di breve raggio. Soltanto i momenti magnetici degli atomi più vicini interagiscono, e in due tipi di situazione compaiono legami frustrati. La prima corrisponde al caso in cui interazioni ferromagnetiche fra atomi adiacenti sono contrastate da interazioni antiferromagnetiche fra atomi vicini, ma posti a distanze maggiori. La seconda compare nei sistemi in cui si ha accoppiamento antiferromagnetico tra i momenti magnetici di atomi adiacenti localizzati nelle posizioni cristallografiche di un reticolo costruito a partire da celle elementari le cui facce possiedono un numero dispari di lati. L'esempio più facile da visualizzare è il reticolo triangolare piano (*si veda l'illustrazione in alto in questa pagina*), in cui i tre momenti magnetici che occupano i vertici di un triangolo equilatero non possono essere simultaneamente a due a due antiparalleli.

Sono dati due esempi di legami frustrati. La frustrazione nasce dalla competizione fra interazioni ferromagnetiche e interazioni antiferromagnetiche. Nell'esempio scelto i momenti magnetici localizzati su ioni adiacenti (che determinano la distanza minima tra ioni dotati di momento magnetico) sono accoppiati tramite interazioni ferromagnetiche J_F che, un po' alla volta, tendono a far allineare tutti i momenti (*a*). Come spesso accade però nei casi reali, i momenti magnetici degli ioni immediatamente più lontani sono accoppiati da interazioni antiferromagnetiche J_{AF} e hanno dunque tendenza a orientarsi in modo antiparallelo. Si osserva che questi legami antiferromagnetici non sono soddisfatti e che è quindi impossibile soddisfare simultaneamente tutti i legami possibili. Un altro tipo di frustrazione deriva dalle interazioni antiferromagnetiche su un reticolo di tipo triangolare. In *b* i momenti degli ioni posti ai vertici di un triangolo sono accoppiati da interazioni antiferromagnetiche J_{AF}. Se i legami tra i momenti magnetici 1, 2 e 2, 3 sono soddisfatti, il legame tra 1 e 3 non potrà esserlo. Anche in questo caso non esiste alcuna configurazione tale che i tre momenti possano essere a due a due antiparalleli, ossia che i tre legami siano soddisfatti simultaneamente.

Gli esempi di composti che presentano le caratteristiche necessarie di disordine e di frustrazione sono molto vari. Oltre alle leghe intermetalliche, si possono citare alcuni tipici isolanti. Il caso più studiato è senza dubbio quello del solfuro di europio e stronzio $Eu_xSr_{1-x}S$. Gli ioni europio portatori dei momenti magnetici e gli ioni stronzio non magnetici occupano in modo aleatorio le posizioni cristallografiche di un reticolo cubico a facce centrate. La competizione fra le interazioni ferromagnetiche, che si hanno fra atomi adiacenti, e quelle antiferromagnetiche, che insorgono fra atomi più lontani, dà luogo a una struttura ferromagnetica in tutti i campioni in cui la concentrazione di ioni europio sia compresa fra il 55 e il 100 per cento, ma al di sotto del 55 per cento i campioni presentano il comportamento caratteristico dei vetri di spin. Il tiospinello di formula chimica $CdCr_{2x}In_{2-2x}S_4$ è un esempio analogo a quello del solfuro di europio, ma lo stato di vetro di spin si manifesta a concentrazioni di ioni magnetici (ioni cromo) molto più elevate (fino a $x = 0{,}85$). Il fluoruro $CsNiFeF_6$ è un altro caso interessante che presenta una frustrazione legata all'esistenza di interazioni antiferromagnetiche in un reticolo la cui cella elementare è un poliedro a facce triangolari. Il disordine, dovuto alla ripartizione aleatoria in questo reticolo degli ioni magnetici nichel e ferro, porta a una configurazione con tre valori diversi d'interazioni antiferromagnetiche. La particolarità in-

Il picco di suscettività magnetica a campo alternato in corrispondenza della temperatura di congelamento magnetico T_g del vetro di spin fa pensare all'esistenza di una transizione di fase. La suscettività magnetica del campione (rapporto fra la variazione di magnetizzazione e la variazione del campo magnetico alternato applicato) è rappresentata in funzione della temperatura per diversi valori di intensità del campo applicato. Le curve presentano un massimo allargato alla temperatura di congelamento magnetico. Ad alta temperatura le curve ottenute per diversi valori di intensità del campo si congiungono e la suscettività varia seguendo la legge di Curie, cioè in modo inversamente proporzionale alla temperatura. La legge di Curie (*curva in colore*) è caratteristica dello stato paramagnetico. La diminuzione della suscettività quando la temperatura decresce al di sotto di T_g indica che i momenti magnetici non possono più seguire le variazioni del campo magnetico applicato. Se il campo applicato diminuisce, i massimi diventano più acuti. La curva estrapolata a campo nullo presenta un picco appuntito (*curva in nero*). In base a questo risultato tutti i momenti magnetici dovrebbero cessare di rispondere alle variazioni del campo magnetico a una temperatura unica e ben definita, comportamento che ricorda quelli cooperativi tipici delle transizioni di fase.

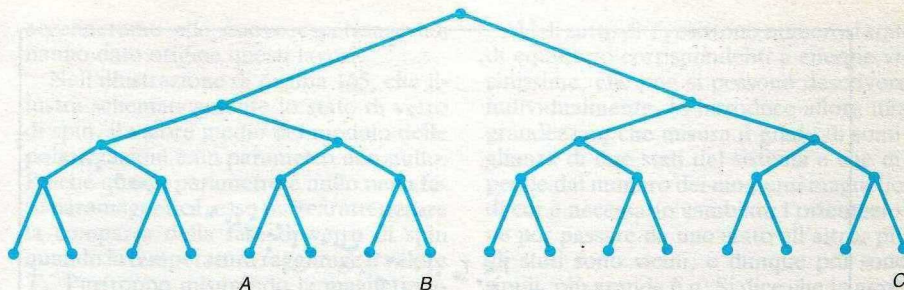
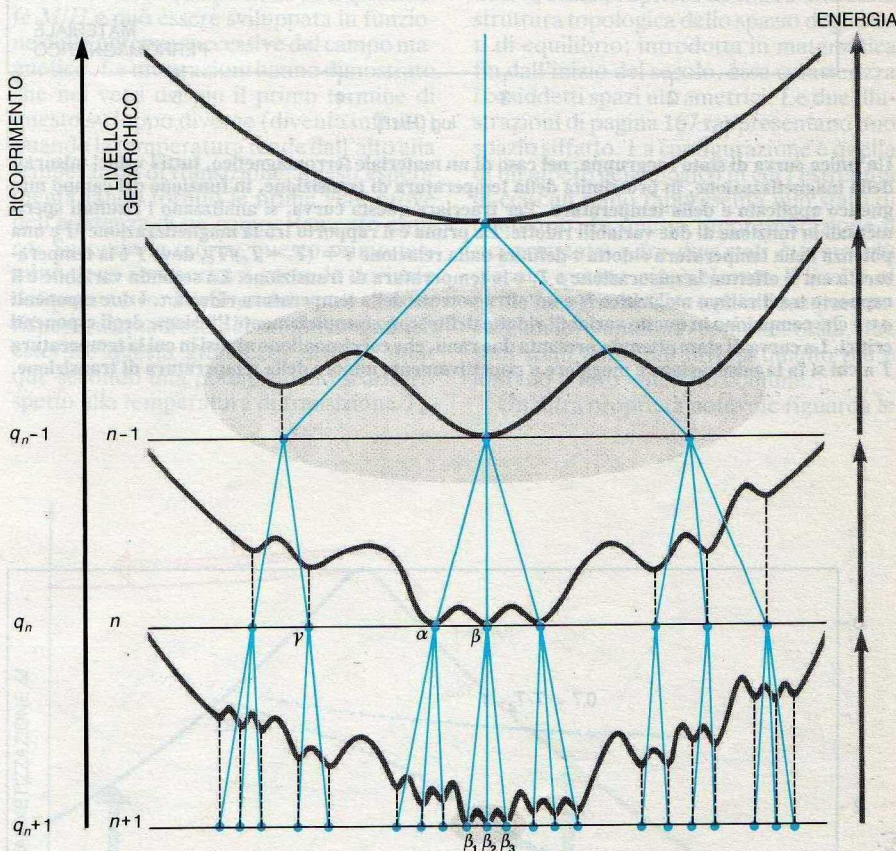
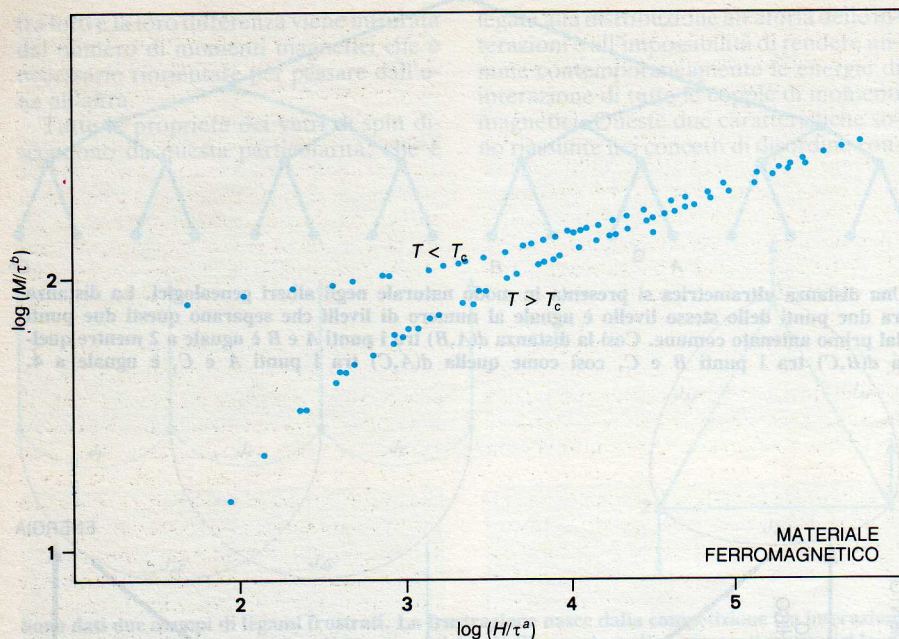
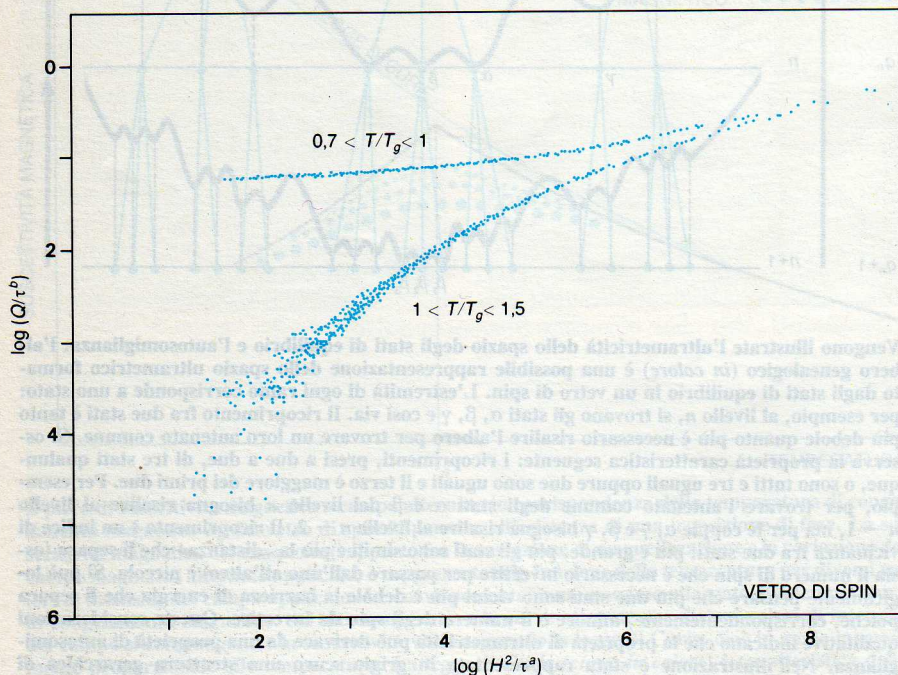
teressante di questo composto è che gli ioni magnetici non sono «diluiti» nel reticolo che occupano. A questi esempi di sostanze isolanti cristalline, nelle quali gli ioni magnetici occupano posizioni ben definite del cristallo, bisogna aggiungere i composti amorfi nei quali le distanze fra ioni adiacenti sono modulate e danno luogo a una distribuzione continua di interazioni. In questa classe di materiali si può ricordare l'alluminosilicato di manganese $(Al_2O_3)_{0,1}(MnO)_{0,5}(SiO_2)_{0,4}$.

Questo elenco, incompleto, dei composti che vengono oggi studiati, fornisce un'idea della generalità del concetto di vetro di spin e della portata di questo campo di ricerca.

Cenni storici

L'interesse per i vetri di spin si è concretizzato a partire dal 1972, quando Vincent D. Cannella e John A. Mydosh effettuarono un esperimento di suscettività magnetica su una lega di oro e ferro. L'interpretazione di questo esperimento rimise in discussione la precedente descrizione delle proprietà dei vetri di spin.

Un campione magnetizzabile collocato in un campo magnetico si magnetizza: la sua suscettività magnetica è il rapporto tra la variazione di magnetizzazione del campione e la variazione del campo applicato che ne è la causa. I risultati sono riportati nell'illustrazione in basso della pagina a fronte, che descrive qualitativamente la suscettività dei vetri di spin immersi in un campo magnetico che inverte periodicamente la propria direzione; questa suscettività, misurata con campi magnetici applicati dell'ordine del centinaio di oersted, presenta un massimo allargato a una temperatura T_g, la «temperatura di congelamento magnetico».

Un comportamento di questo tipo era già stato osservato in diversi sistemi: i fisici l'avevano interpretato invocando gli effetti congiunti delle interazioni fra momenti magnetici e dell'energia di anisotropia (dovuta al fatto che in un cristallo i momenti magnetici prediligono certe direzioni). Le interazioni avrebbero accoppiato certi momenti, che avrebbero così costituito «insiemi» o dominî in grado di rispondere concordemente alle variazioni del campo magnetico alternato. Quando la temperatura è superiore a T_g, la suscettività deriva da un compromesso fra gli effetti opposti del campo magnetico applicato, che tende ad allineare i momenti, e dell'agitazione termica, che tende a sconvolgere l'orientazione: è il regime paramagnetico, in cui la suscettività decresce con il crescere della temperatura, secondo la legge di Curie. In prossimità della temperatura T_g assumerebbero importanza le direzioni di orientazione preferenziali dei momenti magnetici nel cristallo. I dominî rigidi formati dai momenti magnetici accoppiati reagirebbero al campo magnetico applicato con una certa «viscosità» proporzionale al loro volume. Al di sotto della temperatura T_g i dominî non potrebbe-

Una distanza ultrametrica si presenta in modo naturale negli alberi genealogici. La distanza tra due punti dello stesso livello è uguale al numero di livelli che separano questi due punti dal primo antenato comune. Così la distanza $d(A,B)$ tra i punti A e B è uguale a 2 mentre quella $d(B,C)$ tra i punti B e C, così come quella $d(A,C)$ tra i punti A e C, è uguale a 4.

Vengono illustrate l'ultrametricità dello spazio degli stati di equilibrio e l'autosomiglianza: l'albero genealogico (*in colore*) è una possibile rappresentazione dello spazio ultrametrico formato dagli stati di equilibrio in un vetro di spin. L'estremità di ogni ramo corrisponde a uno stato: per esempio, al livello n, si trovano gli stati α, β, γ e così via. Il ricoprimento fra due stati è tanto più debole quanto più è necessario risalire l'albero per trovare un loro antenato comune. Si osserva la proprietà caratteristica seguente: i ricoprimenti, presi a due a due, di tre stati qualunque, o sono tutti e tre uguali oppure due sono uguali e il terzo è maggiore dei primi due. Per esempio, per trovare l'antenato comune degli stati α e β del livello n bisogna risalire al livello $n - 1$, ma per le coppie α, γ e β, γ bisogna risalire al livello $n - 2$. Il ricoprimento è un indice di vicinanza fra due stati: più è grande, più gli stati sono simili e più la «distanza» che li separa (ossia il numero di spin che è necessario invertire per passare dall'uno all'altro) è piccola. Si può logicamente pensare che più due stati sono vicini più è debole la barriera di energia che li separa poiché, corrispondentemente, minore è il numero degli spin da invertire. Queste considerazioni qualitative indicano che la proprietà di ultrametricità può derivare da una proprietà di autosomiglianza. Nell'illustrazione è stata rappresentata in grigio scuro una struttura gerarchica di buche di potenziale corrispondenti agli stati. A ogni livello i si distinguono solo gli stati il cui ricoprimento è inferiore a un valore q_i (cioè la distanza è superiore al valore corrispondente). Così al livello $n + 1$ corrispondente al ricoprimento q_{n+1} gli stati $β_1$, $β_2$ e $β_3$ sono differenziati, mentre al livello n non lo sono. Più la temperatura è bassa più è fine il livello di osservazione (ossia q_i è elevato). Gli stati del livello n sono una microstruttura degli stati dei livelli precedenti. Alla scala n la microstruttura è «simile», previo ingrandimento, alla macrostruttura dei livelli precedenti.

Un'unica curva di stato raggruppa, nel caso di un materiale ferromagnetico, tutti i valori misurati della magnetizzazione, in prossimità della temperatura di transizione, in funzione del campo magnetico applicato e della temperatura. Per tracciare questa curva, si analizzano i risultati sperimentali in funzione di due variabili ridotte. La prima è il rapporto tra la magnetizzazione M e una potenza della temperatura ridotta τ definita dalla relazione $\tau = (T - T_c)/T_c$, dove T è la temperatura a cui si effettua la misurazione e T_c è la temperatura di transizione. La seconda variabile è il rapporto tra il campo magnetico H e un'altra potenza della temperatura ridotta τ. I due esponenti a e b che compaiono in queste variabili ridotte definiscono completamente l'insieme degli esponenti critici. La curva di stato ottenuta presenta due rami, che corrispondono ai casi in cui la temperatura T a cui si fa la misurazione è maggiore o rispettivamente minore della temperatura di transizione.

Nel caso di un vetro di spin si ritrova esattamente lo stesso tipo di funzione di stato. La magnetizzazione M è sostituita da una grandezza Q che ha il comportamento del parametro d'ordine. L'intensità del campo magnetico H è sostituita dal suo quadrato H^2. La grandezza Q corrisponde, al di sopra della temperatura di transizione, alla differenza tra M/H (magnetizzazione misurata divisa per l'intensità del campo) e la suscettività iniziale (dM/dH a campo nullo).

ro più seguire le variazioni troppo rapide del campo e la suscettività tenderebbe dunque a decrescere. Secondo questa interpretazione il massimo osservato dovrebbe essere molto largo, perché non tutti i domini avrebbero la stessa dimensione. Ciascun dominio «congelerebbe» a una temperatura diversa e T_g sarebbe in sostanza la misura del valore medio delle temperature di congelamento magnetico dei singoli domini.

I risultati di Cannella e Mydosh contrastavano con questa interpretazione: nella lega da essi studiata il massimo è tanto più acuto quanto minore è il campo magnetico applicato e la curva estrapolata a campo nullo (*si veda l'illustrazione in basso a pagina 166*) presenta un massimo appuntito corrispondente a una discontinuità. È il picco di suscettività dei vetri di spin. Benché la teoria dei domini di spin descriva correttamente le proprietà di certi materiali, essa non vale nel caso dei vetri di spin. In questi ultimi la temperatura di congelamento magnetico, al di sotto della quale l'insieme dei momenti magnetici cessa di seguire la variazione rapida del campo, è molto netta: il comportamento osservato è caratteristico di un fenomeno cooperativo e la temperatura di congelamento magnetico è una temperatura critica legata a una transizione di fase. Questa ipotesi è stata il vero punto di partenza dello studio sistematico, sia teorico sia sperimentale, dei vetri di spin.

Aspetti statici: la transizione di fase

Il concetto di transizione di fase suscita molto interesse perché descrive un aspetto affatto generale del comportamento della materia. Un dato materiale può trovarsi in stati molto diversi a seconda della temperatura e delle condizioni esterne: un esempio comunissimo è costituito dai tre stati, gassoso, liquido e solido, dell'acqua; il passaggio da uno stato all'altro è una transizione di fase e avviene in maniera brusca.

Ogni materiale è formato da un numero grandissimo di elementi costitutivi identici. Se si conoscono le interazioni fra questi elementi e la statistica alla quale essi obbediscono (cioè la probabilità che questi elementi abbiano una data energia a una temperatura fissata), si possono prevedere le condizioni di esistenza delle diverse fasi come pure i parametri che determinano il passaggio da una fase all'altra.

Le fasi si differenziano in generale per le loro diverse simmetrie. La fase più disordinata, che è anche la più simmetrica (isotropa), si presenta a temperatura elevata. Al procedere del raffreddamento si raggiunge una temperatura critica, a partire dalla quale un parametro fisico del sistema passa da zero a un valore finito; il parametro che caratterizza la nuova fase si chiama parametro d'ordine. La comparsa di un suo valore non nullo abbassa la simmetria del sistema; in effetti le simmetrie che modificano il valore del parametro d'ordine non fanno più parte del

gruppo di simmetria della nuova fase più ordinata: alla temperatura di transizione vi è una rottura di simmetria.

In corrispondenza della temperatura di transizione T_c certe grandezze termodinamiche, come il calore specifico o la suscettività magnetica, hanno un comportamento particolare chiamato singolarità. Quando la temperatura T tende alla temperatura di transizione T_c, queste grandezze, dette critiche, divergono (ossia tendono all'infinito) secondo leggi del tipo $(T - T_c)^{-\gamma}$, dove l'esponente γ non dipende né dai valori delle interazioni fra gli elementi del materiale né dalla configurazione strutturale di quest'ultimo, ma soltanto da due parametri: il numero di dimensioni del materiale e il numero di gradi di libertà dei suoi elementi costitutivi (per esempio una catena di spin ha dimensione uno e ogni spin può avere tre gradi di libertà). Il comportamento critico ha dunque un carattere di universalità e non dipende in alcun modo dalla natura del materiale. Questo è un aspetto importante della teoria delle transizioni di fase.

Per meglio illustrare questa proprietà fondamentale, consideriamo l'esempio tipico della transizione ferromagnetica. Nel corso di questa transizione il sistema passa da uno stato paramagnetico, in cui tutti i momenti magnetici elementari fluttuano e sono, in un dato istante, orientati in modo aleatorio, a uno stato ferromagnetico, in cui la probabilità che i momenti siano orientati secondo una direzione privilegiata dello spazio diviene non nulla e aumenta via via che la temperatura decresce. Il parametro d'ordine è la magnetizzazione M del sistema (la somma vettoriale di tutte le polarizzazioni degli spin): questa grandezza, nulla nella fase paramagnetica, alla temperatura di transizione diviene non nulla. La suscettività magnetica dM/dH a campo applicato H nullo, detta suscettività iniziale, è una delle grandezze termodinamiche caratteristiche che, al tendere della temperatura T a T_c, diverge come una potenza della differenza $T - T_c$. L'esponente critico γ di questa legge dipende solo (nel caso di un campione a tre dimensioni, per esempio) dal numero di gradi di libertà dei momenti magnetici del campione. Quando i singoli momenti possono orientarsi in un'unica direzione dello spazio (ma nei due versi), si ritrova un modello che è stato ben analizzato sotto il profilo teorico, il modello di Ising.

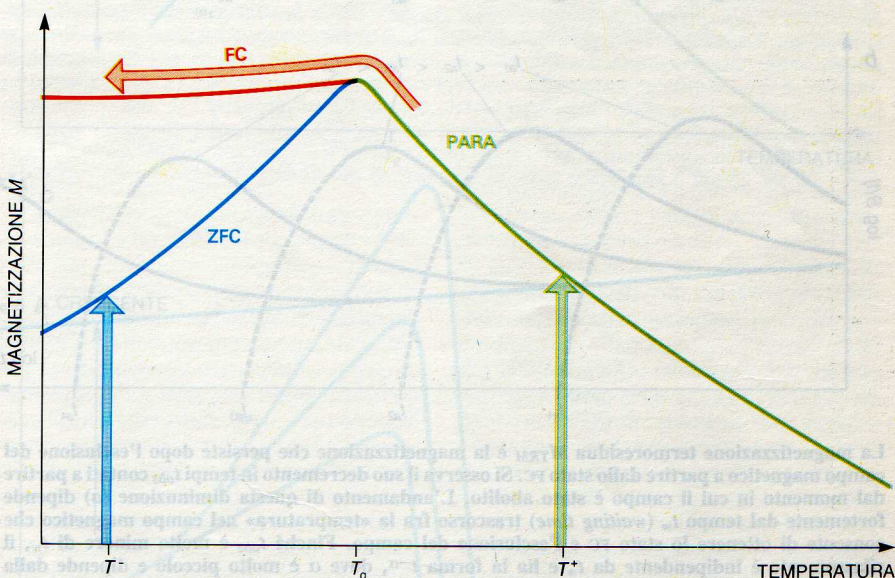
Presentato così, il concetto di transizione di fase permette di capire perché la descrizione del congelamento magnetico del vetro di spin in termini di transizione di fase appaia come un'assurdità. Che tipo di ordine si può definire in un sistema in cui le orientazioni dei momenti magnetici sono assolutamente aleatorie? Che tipo di rottura di simmetria può giustificare l'esistenza di una vera e propria transizione? Qual è il parametro d'ordine?

A queste domande rispondono i numerosi lavori teorici pubblicati dopo la prima osservazione di un picco di suscettività; accenneremo alle nuove esperienze cui hanno dato origine questi lavori.

Nell'illustrazione di pagina 165, che illustra schematicamente lo stato di vetro di spin, il valore medio del modulo delle polarizzazioni è un parametro non nullo. Poiché questo parametro è nullo nella fase paramagnetica, esso può caratterizzare la comparsa della fase di vetro di spin quando la temperatura raggiunge il valore T_g. Purtroppo misurando la magnetizzazione (somma dei vettori di polarizzazione) non si riesce a determinare questo tipo di valore medio, ma i teorici hanno dimostrato che le variazioni di questa media in prossimità della temperatura di transizione sono proporzionali alla cosiddetta suscettività non lineare. Quest'ultima grandezza è la differenza tra la suscettività iniziale dM/dH (a campo nullo) e il quozientе M/H e può essere sviluppata in funzione delle potenze successive del campo magnetico. Le misurazioni hanno dimostrato che nei vetri di spin il primo termine di questo sviluppo diverge (diventa infinito) quando la temperatura tende dall'alto alla temperatura di transizione: questi esperimenti hanno fornito le prime vere prove dell'esistenza di un comportamento critico. Secondo la teoria, questo primo termine deve comportarsi come la derivata del parametro d'ordine del vetro di spin rispetto al quadrato del campo magnetico esterno: questa grandezza diverge dunque secondo una potenza dello scarto rispetto alla temperatura di transizione T_g.

Al di sotto di T_g esistono numerosi stati di equilibrio corrispondenti a energie vicinissime, che non si possono descrivere individualmente. Si introduce allora una grandezza q che misura il grado di somiglianza di due stati del sistema e che dipende dal numero dei momenti magnetici di cui è necessario cambiare l'orientazione per passare da uno stato all'altro: più gli stati sono vicini, e dunque più sono simili, più grande è q. Si dice che la grandezza q misura il «ricoprimento» fra due stati.

Dalle teorie attuali deriva una proprietà notevole dello spazio degli stati di equilibrio: se si scelgono a caso tre stati di equilibrio si ha che tutti e tre i loro ricoprimenti q sono uguali, oppure due sono uguali e il terzo è più grande dei primi due. Questa proprietà definisce una certa struttura topologica dello spazio degli stati di equilibrio; introdotta in matematica fin dall'inizio del secolo, essa caratterizza i cosiddetti spazi ultrametrici. Le due illustrazioni di pagina 167 rappresentano uno spazio siffatto. La configurazione è quella di un albero genealogico in cui l'estremità di ciascun ramo rappresenta uno stato di equilibrio. In questa rappresentazione il ricoprimento q fra due stati è tanto più piccolo quanto più è elevato il livello al quale bisogna risalire nell'albero per trovare un loro antenato comune: ossia, due stati sono tanto più diversi quanto più è lontano il loro antenato comune.

Un'altra proprietà notevole riguarda le

La presenza di un fenomeno di isteresi conferma l'esistenza della fase di vetro di spin. Quando al di sopra della temperatura di transizione T_g (per esempio alla temperatura T^+), si applica un campo H, compare immediatamente una magnetizzazione (*freccia in verde*). Il materiale è nella fase paramagnetica e la sua magnetizzazione segue la legge di Curie-Weiss: varia cioè come l'inverso della temperatura (*curva in verde*). Se, mantenendo il campo applicato, si attua un raffreddamento a partire dallo stato paramagnetico, la magnetizzazione resta costante al di sotto della temperatura T_g (*curva in rosso*) finché non si elimina il campo. La magnetizzazione ottenuta in queste condizioni si chiama FC (da *field cooled*). Se il campione viene prima raffreddato al di sotto di T_g, a campo nullo, l'applicazione del campo (per esempio alla temperatura T^-) è seguita immediatamente dalla comparsa di una magnetizzazione inferiore alla magnetizzazione FC (*freccia in blu*). La magnetizzazione ottenuta, che si chiama ZFC (da *zero field cooled*), non è stabile e subisce un rilassamento verso la magnetizzazione FC per tutto il tempo in cui il campo resta applicato.

energie associate ai diversi stati. Quando si considerano le energie delle configurazioni magnetiche, gli stati di equilibrio corrispondono a minimi. Osservando però da vicino il fondo di un minimo allargato si vedono comparire numerosi minimi più piccoli. Quando si osserva uno qualunque di questi piccoli minimi aumentando ancora la definizione, si scopre una nuova struttura fine di questo minimo, che è analoga a quella di partenza. Questa proprietà di autosomiglianza dei minimi di energia è ancora soltanto un'ipotesi, ma è fortemente sostenuta dal comportamento dei vetri di spin (*si veda l'illustrazione a pagina 167 in basso*).

Qual è ora il parametro d'ordine? In queste teorie esso non è una grandezza macroscopica, bensì una funzione che dipende dalla distribuzione dei ricoprimenti fra stati di equilibrio. Questo parametro d'ordine non può essere misurato direttamente, tuttavia è possibile studiare per via sperimentale una grandezza scalare Q che lo caratterizza. Q è legata alla magnetizzazione del sistema, rilevata in funzione della temperatura quando a una temperatura superiore a T_g viene applicato preliminarmente un campo magnetico continuo H (*si veda l'illustrazione a pagina 169*). Al di sopra della temperatura T_g, Q è equivalente alla suscettività non lineare il cui comportamento è già stato descritto, mentre al di sotto di T_g, Q corrisponde alla differenza fra la magnetizzazione e l'estrapolazione a bassa temperatura della magnetizzazione lineare. Nei vetri di spin Q ha lo stesso ruolo che ha la magnetizzazione M in un corpo ferromagnetico. I teorici hanno peraltro dimostrato che, nei vetri di spin, il quadrato del campo magnetico è l'equivalente del campo in una sostanza ferromagnetica.

L'esistenza di una transizione di fase ha potuto allora essere confermata mettendo in evidenza la legge di scala che descrive il comportamento di Q intorno alla temperatura di congelamento magnetico. Questa legge stabilisce che esiste una funzione unica (funzione di scala) che lega le due seguenti variabili ridotte: la prima corrispondente al rapporto fra Q e una potenza della temperatura ridotta $(T - T_g)/T_g$; la seconda uguale al rapporto fra il quadrato del campo magnetico e un'altra potenza della temperatura ridotta. I due esponenti vengono valutati mediante verifica dell'unicità della funzione di scala ottenuta tracciando, in funzione delle variabili ridotte, tutte le curve di magnetizzazione rilevate in funzione del campo magnetico e della temperatura in prossimità di T_g. Dai due esponenti così definiti si ricava l'insieme degli esponenti critici del sistema. Le illustrazioni di pagina 168 mostrano la funzione di scala tracciata nel caso di un vetro di spin isolante, confrontata con quella di un materiale ferromagnetico. È abbastanza interessante osservare la somiglianza fra le curve e poter interpretare in maniera così fine i risultati ottenuti in sistemi disordinati che, in un primo momento, appaiono di una complessità estrema.

Gli esponenti critici sono stati misurati per diversi composti. Nel caso dei composti intermetallici, i valori trovati possono essere ripartiti in due classi di universalità come nel caso dei sistemi tridimensionali ordinati. Attualmente la situazione è meno chiara nel caso degli isolanti, per i quali i valori degli esponenti sono relativamente disparati.

La magnetizzazione termoresidua M_{TRM} è la magnetizzazione che persiste dopo l'esclusione del campo magnetico a partire dallo stato FC. Si osserva il suo decremento in tempi t_{obs} contati a partire dal momento in cui il campo è stato abolito. L'andamento di questa diminuzione (*a*) dipende fortemente dal tempo t_w (*waiting time*) trascorso fra la «temprature» nel campo magnetico che consente di ottenere lo stato FC e l'esclusione del campo. Finché t_{obs} è molto minore di t_w, il rilassamento è indipendente da t_w e ha la forma $t^{-\alpha}$, dove α è molto piccolo e dipende dalla temperatura. Poi le curve abbandonano il tratto comune, e ciò tanto più tardi quanto maggiore è t_w. Il rilassamento invariante per t_{obs} molto minore di t_w è dunque il rilassamento corrispondente all'equilibrio termodinamico, poiché è l'unico che si conserva quando t_w tende all'infinito. Il parametro temporale che in effetti condiziona questo invecchiamento del processo di rilassamento è l'età del sistema, t_a, pari a $t_w + t_{obs}$. Per età determinate si può tracciare, a partire dalla pendenza logaritmica del rilassamento $S(t)$, la distribuzione dei tempi di rilassamento del sistema alle età t_a successive (*b*). La forma della distribuzione resta invariante sul diagramma logaritmico e la porzione in cui t è molto minore di t_a rappresenta la distribuzione all'equilibrio. Si può far subire alle curve che rappresentano $M_{TRM}(t)$ la trasformazione di scala che corrisponde alla conservazione di un'età costante $t_a = t_w$. Si ottengono curve (*tratteggiate in a*) che rappresentano l'evoluzione di un sistema descritto dalla distribuzione dei tempi di rilassamento corrispondente all'età t_w. Tutte le curve sono identiche, e appaiono solo traslate le une rispetto alle altre; esse corrispondono a un rilassamento di Kohlrausch $\exp - (t/t_0)^{\beta}$, dove t_0 e β dipendono dalla temperatura.

Aspetti dinamici

Le proprietà dinamiche dei vetri di spin sono molto più vicine a quelle dei sistemi senza transizione di fase, come i vetri classici o i polimeri, che non a quelle delle fasi magnetiche classiche. I vetri di spin sono caratterizzati da tempi di rilassamento (tempi necessari al ripristino dell'equilibrio dopo una modificazione delle condizioni esterne) estremamente lunghi e da effetti d'invecchiamento che mascherano i rilassamenti intrinseci.

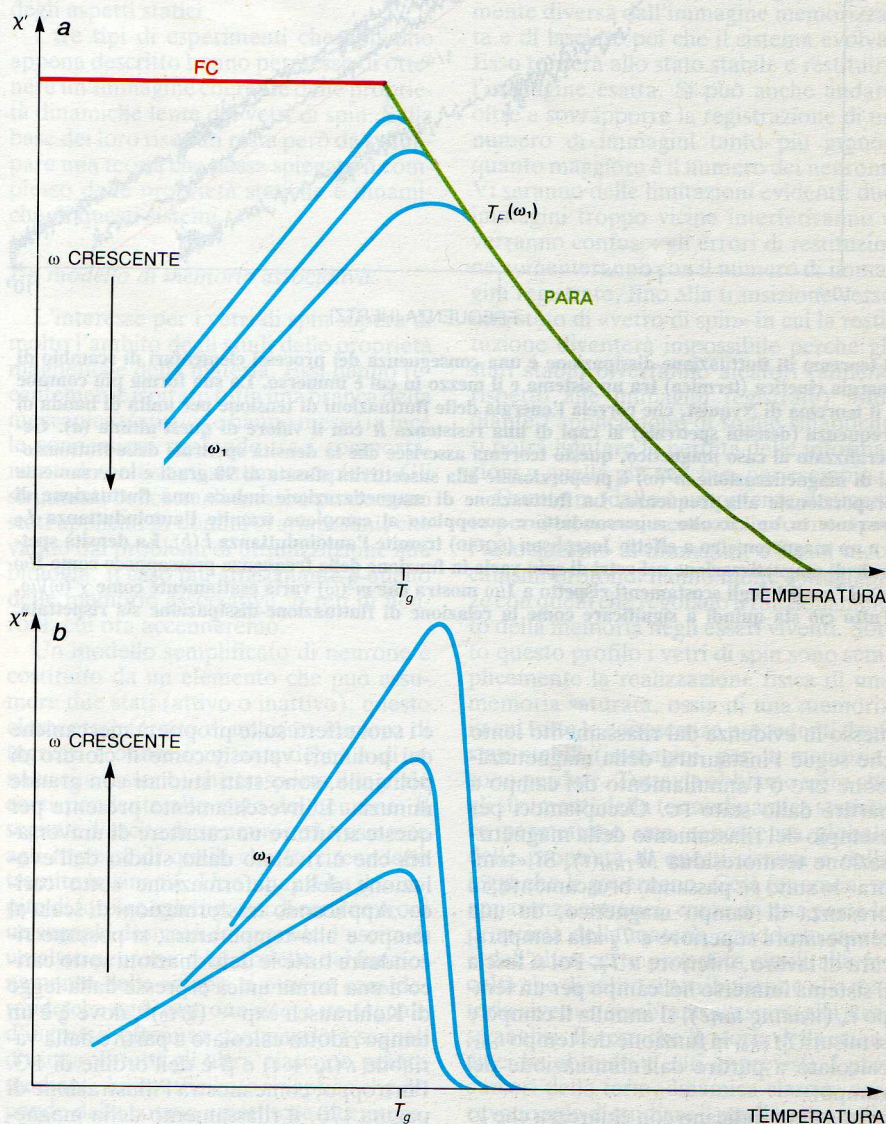
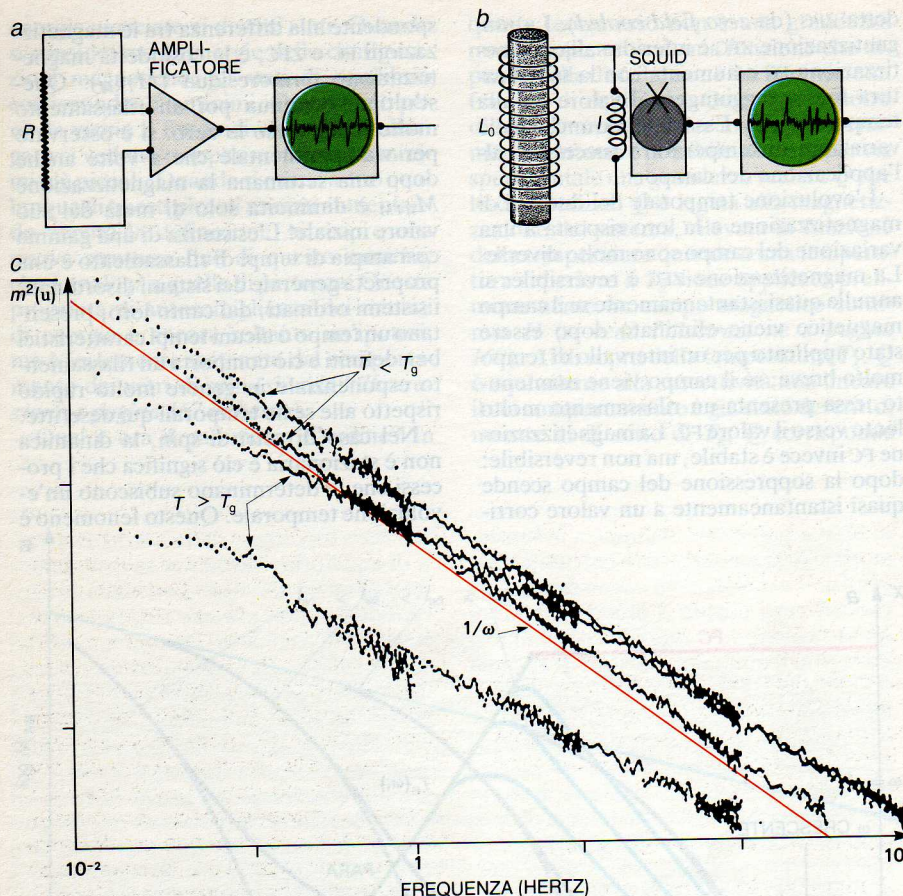
La dinamica di un sistema è l'insieme delle leggi che ne reggono l'evoluzione temporale. Nel corso di questa evoluzione il sistema tende a occupare stati stabili di energia libera minima. Questo rilassamento è semplice quando vi è un solo minimo di energia: qualunque sia la configurazione di partenza il sistema evolve sempre verso lo stesso stato stabile, che lo definisce completamente. Ciò accade

in particolare nella fase paramagnetica, quando la temperatura del sistema è superiore a T_g. Se invece esistono più minimi e se per passare dall'uno all'altro è necessario attraversare configurazioni di energia elevata (cioè superare barriere di energia), lo stato stabile di arrivo dipende dalla configurazione di partenza. Il sistema si dice non ergodico se esistono barriere di energia che non possono essere superate al limite neppure in un tempo infinito. Quando i minimi si ricavano l'uno dall'altro mediante operazioni di simmetria note, si è ricondotti al problema di un solo minimo; è questo il caso dei mezzi ordinati classici. Abbiamo visto che nei sistemi complessi come i vetri di spin esiste un'infinità di minimi il cui numero non può essere ridotto da alcuna simmetria di struttura e, in corrispondenza, vi è una distribuzione quasi continua di picchi che fungono da barriere fra questi minimi (*si veda l'illustrazione a pagina 167 in basso*). Poiché il tempo medio necessario perché il sistema superi una data barriera cresce esponenzialmente con l'altezza di quest'ultima, l'evoluzione del sistema è retta da una distribuzione molto ampia dei tempi di rilassamento. Il suo comportamento dinamico si estende su tempi lunghissimi. I tempi di rilassamento che si possono misurare in laboratorio costituiscono solo una «finestra» ridottissima rispetto all'ampiezza della distribuzione di questi tempi.

Rilassamento in tempi lunghi

Una delle prime manifestazioni delle proprietà dinamiche della fase di vetro di spin è la presenza, al di sotto della temperatura di congelamento magnetico, di un'isteresi molto cospicua. In una fase magnetica classica (che comprende un solo dominio), l'applicazione di un campo magnetico produce istantaneamente una magnetizzazione (in tempi dell'ordine di 10^{-11} secondi), che resta costante, a una data temperatura, finché il campo rimane applicato e sparisce con la stessa rapidità quando il campo viene soppresso. Nella fase di vetro di spin la magnetizzazione corrispondente a un campo assegnato dipende dalla storia termica del materiale prima e dopo l'applicazione del campo (*si veda l'illustrazione a pagina 169*). Quando la magnetizzazione viene prodotta al di sopra della temperatura di congelamento magnetico, nella fase paramagnetica, e il materiale viene poi raffreddato al di sotto di T_g, la magnetizzazione è pressappoco costante in tutta la fase a bassa temperatura e dipende poco dall'evoluzione termica successiva: a una data temperatura ha sempre lo stesso valore, quali che siano le variazioni termiche subite fra due misurazioni. Questa magnetizzazione viene chiamata magnetizzazione FC (da *field cooled*). Se il materiale è stato raffreddato al di sotto della temperatura di congelamento magnetico in un campo nullo, la successiva applicazione del campo magnetico comporta una magnetizzazione detta ZFC (da *zero field cooled*). La magnetizzazione ZFC è inferiore alla magnetizzazione FC e aumenta con la temperatura fino a raggiungere il valore FC alla temperatura T_g. Essa dipende anche dalle variazioni di temperatura successive all'applicazione del campo.

L'evoluzione temporale dei due tipi di magnetizzazione e la loro risposta a una variazione del campo sono molto diverse. La magnetizzazione ZFC è reversibile: si annulla quasi istantaneamente se il campo magnetico viene eliminato dopo essere stato applicato per un intervallo di tempo molto breve; se il campo viene mantenuto, essa presenta un rilassamento molto lento verso il valore FC. La magnetizzazione FC invece è stabile, ma non reversibile: dopo la soppressione del campo scende quasi istantaneamente a un valore corrispondente alla differenza tra le magnetizzazioni FC e ZFC; è la cosiddetta magnetizzazione termoresidua (M_{TRM}). Quest'ultima continua poi un rilassamento molto lento verso lo zero: si è osservato per via sperimentale che a volte anche dopo una settimana la magnetizzazione M_{RTM} è diminuita solo di metà del suo valore iniziale! L'esistenza di una gamma così ampia di tempi di rilassamento è una proprietà generale dei sistemi disordinati; i sistemi ordinati, dal canto loro, presentano un tempo o alcuni tempi caratteristici ben definiti e ciò comporta un rilassamento esponenziale in genere molto rapido rispetto alle scale temporali qui descritte.

Nel caso di vetri di spin, la dinamica non è stazionaria e ciò significa che i processi che la determinano subiscono un'evoluzione temporale. Questo fenomeno è

La risposta (magnetizzazione) a un campo magnetico di frequenza ω dei diversi sistemi di vetri di spin varia per effetto di certe caratteristiche dei materiali studiati, ma si può schematicamente individuare un comportamento comune. Ad alta temperatura il materiale è paramagnetico: la suscettività segue la legge di Curie. Esso abbandona questo comportamento a una temperatura $T_f(\omega)$ che diminuisce al diminuire di ω. Ciò si traduce nella presenza di un massimo di suscettività in fase $\chi'(\omega)$ a $T_f(\omega)$ (*a*) e nella comparsa di una suscettività sfasata di 90 gradi $\chi''(\omega)$ (*b*) non nulla al di sotto di $T_f(\omega)$, che risale rapidamente a $T_f(\omega)$. La suscettività χ'' corrisponde alla dissipazione.

Il teorema di fluttuazione-dissipazione è una conseguenza dei processi elementari di scambio di energia cinetica (termica) tra un sistema e il mezzo in cui è immerso. La sua forma più comune è il teorema di Nyquist, che correla l'energia delle fluttuazioni di tensione per unità di banda di frequenza (densità spettrale) ai capi di una resistenza R con il valore di quest'ultima (*a*). Generalizzato al caso magnetico, questo teorema asserisce che la densità spettrale delle fluttuazioni di magnetizzazione $\overline{m^2(\omega)}$ è proporzionale alla suscettività sfasata di 90 gradi e inversamente proporzionale alla frequenza. La fluttuazione di magnetizzazione induce una fluttuazione di corrente in un circuito superconduttore accoppiato al campione tramite l'autoinduttanza L_0 e a un magnetometro a effetto Josephson (SQUID) tramite l'autoinduttanza l (*b*). La densità spettrale di magnetizzazione nei vetri di spin varia in funzione della frequenza pressappoco come $1/\omega$ (*c*). L'analisi degli scostamenti rispetto a $1/\omega$ mostra che $\overline{m^2(\omega)}$ varia esattamente come $\chi''(\omega)/\omega$. Tutto ciò sta quindi a significare come la relazione di fluttuazione-dissipazione sia rispettata.

messo in evidenza dal rilassamento lento che segue l'instaurarsi della magnetizzazione ZFC o l'annullamento del campo a partire dallo stato FC. Occupiamoci per esempio del rilassamento della magnetizzazione termoresidua $M_{TRM}(t)$. Si «tempra» lo stato FC passando bruscamente, in presenza di campo magnetico, da una temperatura superiore a T_g alla temperatura di lavoro, inferiore a T_g. Poi si lascia il sistema immerso nel campo per un tempo t_w (*waiting time*), si annulla il campo e si misura M_{TRM} in funzione del tempo t_{obs}, calcolato a partire dall'eliminazione del campo.

I risultati indicano con chiarezza che lo stato FC non è uno stato di equilibrio poiché durante il tempo trascorso in questo stato i processi che determinano la dinamica si evolvono, dando luogo a un rilassamento che dipende da t_w. Questo fenomeno, detto d'invecchiamento, è comune a un gran numero di strutture disordinate e i suoi effetti sulle proprietà meccaniche dei polimeri vetrosi, come il cloruro di polivinile, sono stati studiati con grande minuzia. L'invecchiamento presenta per queste strutture un carattere di universalità che è rivelato dallo studio dell'evoluzione della deformazione sotto carico. Applicando trasformazioni di scala al tempo e alla temperatura, si possono ricondurre tutte le deformazioni sotto carico a una forma unica espressa dalla legge di Kohlrausch $\exp - (\xi/\tau_0)^\beta$, dove ξ è un tempo ridotto calcolato a partire dalla variabile $t/(t_w + t)$ e β è dell'ordine di 1/3. Purtroppo, come mostra l'illustrazione di pagina 170, il rilassamento della magnetizzazione termoresidua nei vetri di spin non possiede questo carattere di universalità, ma può essere descritto dal prodotto di due tipi di risposte, una stazionaria (indipendente dal tempo t_w) e una (dipendente da t_w) che rappresenta l'invecchiamento. In generale il rilassamento stazionario è rappresentato correttamente da una potenza del tempo, $t^{-\alpha}$, dove α è un esponente piccolo e dipendente dalla temperatura. Il rilassamento d'invecchiamento, preponderante quando t_{obs} è maggiore di t_w, dipende da t_w secondo una legge di scala temporale identica a quella dei polimeri vetrosi; anche la sua forma è quella della legge di Kohlrausch, ma l'esponente β varia con la temperatura.

La risposta sinusoidale

Un altro esperimento importante è la misurazione della risposta a un'eccitazione sinusoidale pura. In un sistema lineare (un sistema è lineare se la sua risposta alla somma di due eccitazioni è uguale alla somma delle singole risposte), questa risposta è direttamente collegata al rilassamento di cui abbiamo parlato in precedenza: siccome è possibile scomporre sotto il profilo matematico qualunque evoluzione temporale in una somma di contributi elementari sinusoidali puri (operazione nota col nome di trasformata di Fourier), se si conosce la risposta di un sistema lineare a tutte le frequenze pure si può calcolare la sua risposta a tutte le forme di eccitazione. Si può tra l'altro ritrovare la legge di rilassamento della magnetizzazione termoresidua.

Si sottopone il materiale a un campo che varia nel tempo in modo sinusoidale e si analizza la magnetizzazione risultante. Se il sistema è lineare, anche la sua magnetizzazione è una sinusoide pura con lo stesso periodo, ma la sfasatura è tanto più grande quanto maggiore è il tempo di rilassamento proprio del sistema rispetto a questo periodo. Si distinguono due parti della risposta, una, χ', in fase e l'altra, χ'', sfasata di 90 gradi rispetto al campo. L'illustrazione di pagina 171 mostra un comportamento tipico dei vetri di spin. All'inizio la suscettività in fase χ' presenta un picco in funzione della temperatura. Al diminuire della frequenza questo picco si sposta verso le basse temperature e diventa via via più acuto. Nella fase di vetro di spin la suscettività χ' è tanto più debole quanto più grande è la frequenza; compare però la suscettività sfasata di 90 gradi χ'', che è tanto maggiore quanto più elevata è la frequenza. Ciò dimostra la viscosità della fase di vetro di spin, la cui dissipazione è tanto maggiore quanto più elevata è la frequenza. Le variazioni di χ' e χ'' in funzione della frequenza sono molto diverse da quelle dei sistemi classici, come si può prevedere dall'esistenza, dedotta dalle misure di rilassamento, di un'ampia distribuzione dei tempi caratteristici.

Fluttuazioni libere

Di recente l'arsenale delle misurazioni a bassa frequenza del comportamento di un sistema nella fase di vetro di spin si è arricchito di una nuova tecnica di grandissima potenza. Si tratta dello studio delle fluttuazioni libere della magnetizzazione, in cui si osserva il materiale all'equilibrio

senza alcun tipo di eccitazione. Con questa tecnica si misura il «rumore magnetico» spontaneo del campione.

Ogni esperto di elettronica sa prevedere l'entità del rumore termico di un dispositivo elettronico applicando il famoso teorema di Nyquist, che lega la fluttuazione della tensione ai capi di un elemento elettrico passivo al valore della sua resistenza. Storicamente il teorema di Nyquist è stato la prima enunciazione di una relazione molto generale della termodinamica statistica, chiamata teorema di fluttuazione-dissipazione. (Un sistema all'equilibrio non è immutabile ed evolve continuamente intorno al suo stato di equilibrio dando luogo a fluttuazioni.) Il teorema stabilisce una relazione tra le fluttuazioni di un sistema in equilibrio termodinamico e la sua viscosità. In effetti queste due grandezze hanno la proprietà di essere entrambe conseguenza di interazioni aleatorie fra il sistema studiato e il mezzo in cui è immerso (per esempio il sistema degli spin è soggetto alle vibrazioni termiche del reticolo atomico). In un sistema magnetico la viscosità è il rapporto tra la suscettività sfasata di 90 gradi χ'' e la frequenza. Il teorema di fluttuazione-dissipazione si può evidentemente applicare solo a sistemi ergodici, perché gli insiemi delle configurazioni percorse durante le fluttuazioni libere e durante la risposta a un'eccitazione (misurazione di χ'') devono essere gli stessi. Inoltre esso assume la forma semplice ora enunciata soltanto nei sistemi a dinamica lineare.

Le misurazioni delle fluttuazioni magnetiche sono delicatissime e sono possibili solo da quando sono state introdotte le moderne tecniche di magnetometria a effetto Josephson, le quali permettono di rilevare campi magnetici dell'ordine di 10^{-11} gauss (il campo magnetico terrestre vale circa 0,5 gauss). Queste misurazioni sono molto adatte allo studio dei sistemi disordinati in cui la suscettività sfasata di 90 gradi è in genere rilevante a bassa frequenza. In effetti se la suscettività χ'' è indipendente dalla frequenza (caso limite di una distribuzione uniforme dei tempi di rilassamento), il rumore corrispondente varia come l'inverso della frequenza e quindi a bassa frequenza i contributi sarebbero rilevanti.

L'illustrazione della pagina a fronte fornisce un esempio tipico di spettri di magnetizzazione rilevati in vetri di spin all'equilibrio e dopo un tempo d'invecchiamento sufficiente perché le fluttuazioni siano stazionarie in tutta la gamma di frequenze. Vi è perfetto accordo con i valori dedotti, sfruttando la relazione di fluttuazione-dissipazione, dai risultati delle misurazioni della suscettività sfasata di 90 gradi χ''. Ciò può sembrare sorprendente, visto che conosciamo il carattere non ergodico dei vetri di spin che rivelano differenze di comportamento a seconda che la loro «tempratura» sia stata effettuata in presenza di campo magnetico o a campo nullo. Come può la relazione di fluttuazione-dissipazione descrivere il comportamento di materiali di questo genere? Si possono soltanto avanzare alcune ipotesi. Le due misurazioni a cui si riferisce la relazione di fluttuazione-dissipazione sono misurazioni di frequenze che obbediscono agli stessi vincoli temporali: una misurazione effettuata a una data frequenza ha una durata caratteristica, che è l'inverso di questa frequenza. Si può allora pensare che gli insiemi di configurazioni esplorati in misurazioni diverse con scale temporali equivalenti abbiano, benché disgiunti, proprietà locali simili. Ciò potrebbe essere dovuto alla proprietà già ricordata dell'autosomiglianza dello spazio degli stati (*si veda l'illustrazione a pagina 167 in basso*), proprietà affatto compatibile con l'ultrametricità che deriva dalle teorie del campo medio ricordate nella discussione degli aspetti statici.

I tre tipi di esperimenti che abbiamo appena descritto hanno permesso di ottenere un'immagine coerente delle proprietà dinamiche lente dei vetri di spin. Sulla base dei loro risultati resta però da sviluppare una teoria che possa spiegare il complesso delle proprietà statiche e dinamiche di questi sistemi.

Un modello di memoria associativa

L'interesse per i vetri di spin supera di molto l'ambito degli studi delle proprietà magnetiche della materia. Essi hanno recentemente ispirato tutta una branca della meccanica statistica e hanno messo in luce le connessioni profonde tra i comportamenti di moltissimi sistemi complessi. Gli strumenti concettuali messi a punto sono stati applicati in dominî svariatissimi, che vanno dai problemi di ottimizzazione alla biologia. Il caso più affascinante è quello delle memorie associative, o reti di neuroni, cui ora accenneremo.

Un modello semplificato di neurone è costituito da un elemento che può assumere due stati (attivo o inattivo); questo elemento è fornito di un certo numero di ingressi (connessioni sinaptiche) e di una uscita (assone) che trasporta il segnale corrispondente allo stato del neurone. Il segnale d'uscita di un neurone è trasmesso agli ingressi di quelli che lo circondano tramite le sinapsi. Una sinapsi ha il compito di fornire agli ingressi di certi neuroni un segnale derivato dal segnale d'uscita di un altro neurone. Un modello del funzionamento delle sinapsi si ottiene considerando che ogni neurone riceve un segnale d'ingresso ottenuto sommando i segnali d'uscita di tutti gli altri, ciascuno pesato mediante un coefficiente sinaptico che dipende dalla coppia emettitore-ricevitore. Un neurone è attivo o inattivo a seconda che la somma dei segnali sinaptici superi o no una soglia caratteristica. È evidente la somiglianza con un modello di Ising in cui ogni spin può avere due orientazioni e le interazioni fra due spin hanno valori distribuiti in modo aleatorio. Abbiamo visto che un sistema del genere possiede stati stabili verso i quali si evolve spontaneamente a partire da qualsiasi configurazione vicina.

Sulla base di questo modello si può costruire una memoria elettronica. L'apprendimento si compie non assegnando a un certo numero di elementi della memoria un valore corrispondente all'informazione da registrare, bensì dando a ciascun coefficiente sinaptico un valore calcolato a partire dall'informazione (l'immagine) da memorizzare. Quindi lo stato stabile dell'intera memoria corrisponde all'informazione esatta, ma, a differenza di una memoria indirizzabile, la memoria associativa ha una dinamica di riconoscimento dell'informazione. Supponiamo di imporre a tutti i neuroni un'immagine leggermente diversa dall'immagine memorizzata e di lasciare poi che il sistema evolva. Esso tornerà allo stato stabile e restituirà l'immagine esatta. Si può anche andare oltre e sovrapporre la registrazione di un numero di immagini tanto più grande quanto maggiore è il numero dei neuroni. Vi saranno delle limitazioni evidenti: due immagini troppo vicine interferiranno e verranno confuse; gli errori di restituzione aumenteranno con il numero di immagini registrate, fino alla transizione verso uno stato di «vetro di spin» in cui la restituzione diventerà impossibile perché gli stati stabili diverranno aleatori (almeno rispetto alle immagini registrate). Mediante l'applicazione di soglie sinaptiche si potrà limitare la perdita delle informazioni a quelle più vecchie, per esempio.

Questo modello è un potente strumento per il riconoscimento delle forme e per l'associazione di immagini, e i suoi meccanismi profondi hanno molte somiglianze con quelli che regolano il funzionamento della memoria negli esseri viventi. Sotto questo profilo i vetri di spin sono semplicemente la realizzazione fisica di una memoria saturata, ossia di una memoria in cui tutte le conoscenze precedenti degli stati stabili (immagini, per la memoria) sono perdute. Tuttavia abbiamo visto che gli strumenti della termodinamica statistica permettono una conoscenza globale delle proprietà di questi sistemi e delle leggi che li governano. Così fenomeni in apparenza estranei, come da una parte le proprietà della materia cosiddetta inerte e dall'altra certe proprietà simili alle facoltà che di solito si attribuiscono alla materia vivente, sembrano prestarsi alla stessa analisi. È vero che la storia della scienza ci ha abituati a esiti imprevisti. I precursori della termodinamica classica cercavano solo di studiare il rendimento delle macchine termiche; e invece gettarono le basi di una scienza i cui principi si applicano a tutti i fenomeni naturali, compresa la vita.

«Vetri di spin»

I modelli matematici di questi materiali, caratterizzati a livello atomico da interazioni magnetiche disordinate, servono per l'analisi di problemi complessi in campi quali la scienza dei calcolatori e la neurobiologia

di Daniel L. Stein

La sporcizia può essere spinta in un angolo o nascosta sotto il tappeto, ma prima o poi richiede attenzione. In fisica «sporcizia» può indicare disordine in una struttura, impurezze in un materiale o competizione fra più interazioni. La sporcizia è nemica dell'ordine. Quantità sufficienti di casualità, di imperfezioni e di discordanze possono distruggere le simmetrie intrinseche che semplificano drasticamente le descrizioni fisiche. Per la maggior parte della storia della fisica la sporcizia è stata spinta in un angolo e gli scienziati si sono dedicati allo studio di sistemi ordinati, quali i cristalli perfetti. Tuttavia, all'inizio degli anni settanta, essi furono costretti a fronteggiare il disordine; la sporcizia che si era depositata nei corridoi della scienza cominciò a venire alla luce. Inizialmente, per esempio, venne prodotto un po' di disordine in un cristallo perfetto per acquisire conoscenze sui vetri, i cui atomi sono «congelati» in posizioni casuali nello spazio; ma questi tentativi fallirono. Escludere quantità troppo grandi di disordine da sistemi per loro natura «sporchi» è come cercare di studiare una pozzanghera di fango pulita.

Uno dei tentativi di maggior successo per comprendere i sistemi disordinati è stato lo studio dei cosiddetti «vetri di spin». La composizione di questo materiale non ha nulla di notevole - qualche atomo di ferro sparso in un reticolo di atomi di rame - ma le sue proprietà magnetiche sono stupefacentemente complesse e talvolta risultano imprevedibili in maniera frustrante. Lo «spin» al quale ci si riferisce è quello quantomeccanico che dà origine a effetti magnetici; con il termine «vetro» si indica invece il disordine nell'orientazione e nelle interazioni degli spin. Il vetro di spin è un sistema sporco per eccellenza.

Le tecniche messe a punto per esaminare i vetri di spin sono state applicate allo studio di problemi complessi in campi estremamente diversificati, quali la scienza dei calcolatori, la neurologia, la biochimica e lo studio dell'evoluzione biologica.

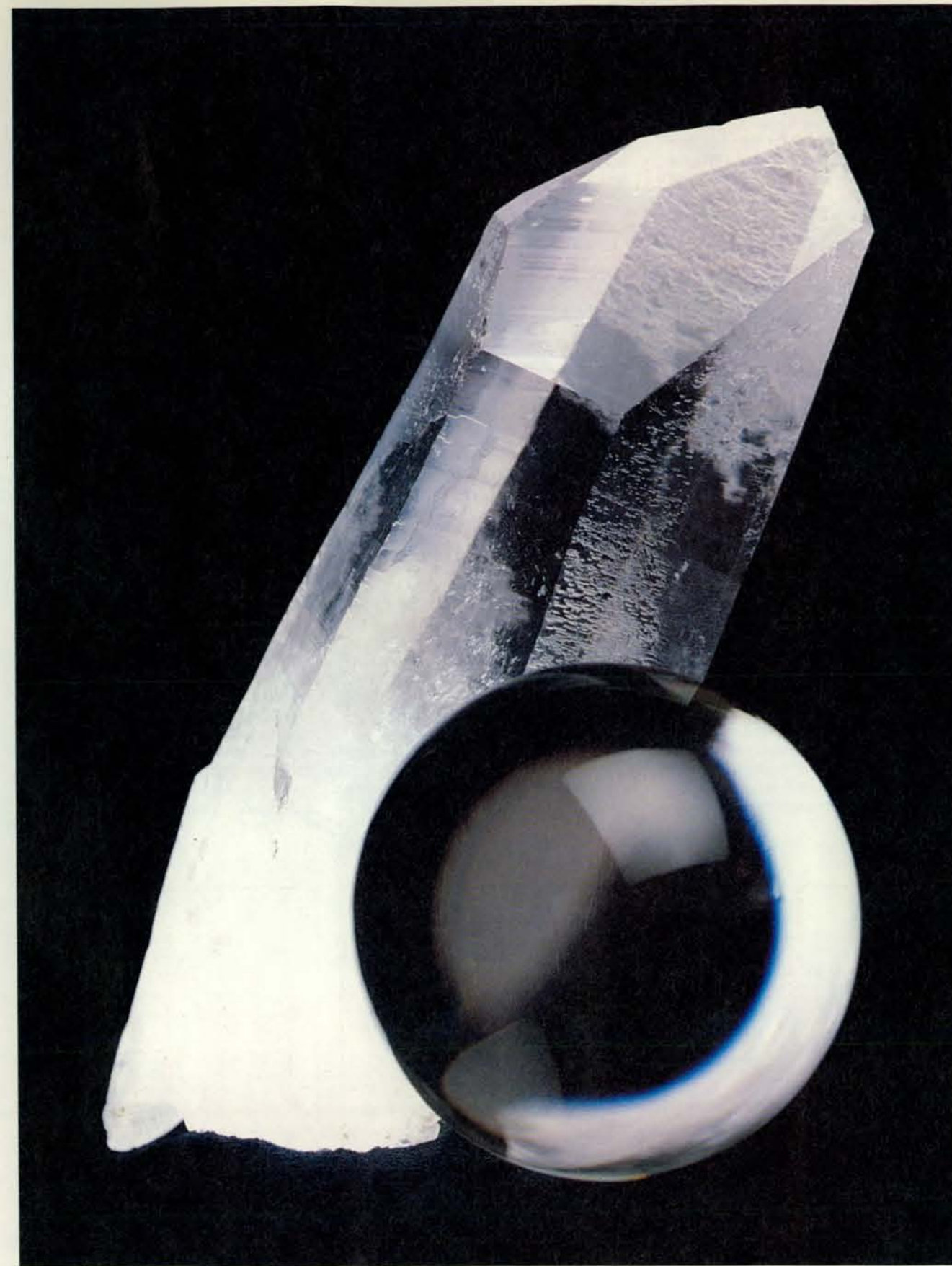
Le interessanti caratteristiche dei vetri di spin, la loro dinamica e la loro complessità sono tutte dovute alle interazioni magnetiche tra gli atomi che li compongono. Certi atomi si comportano come piccolissime barrette magnetiche: generano campi magnetici e ne sono soggetti. La direzione, il verso e l'intensità degli effetti magnetici si possono descrivere mediante una grandezza vettoriale, il momento magnetico. Se un blocco di materiale i cui atomi si comportano come magneti viene esposto a un campo esterno, i momenti magnetici tendono ad allinearsi in una direzione particolare. In alcuni materiali questo allineamento dei momenti magnetici può verificarsi anche a causa di forti effetti interni associati alla struttura atomica.

Uno di questi effetti determina l'allineamento dei momenti magnetici, il quale spiega le forti proprietà magnetiche del ferro, e per questo motivo viene chiamato ferromagnetismo, anche se è riscontrabile in molti altri materiali quali il cobalto e il nichel. Il ferromagnetismo è dovuto alle proprietà quantomeccaniche degli elettroni più interni di questi metalli, proprietà che rendono favorevole dal punto di vista energetico la disposizione parallela dei momenti magnetici di atomi adiacenti. In altre parole, se i momenti magnetici di due atomi ferromagnetici adiacenti hanno lo stesso verso, è necessario fornire energia per orientare uno dei momenti magnetici nel verso opposto; viceversa, se i momenti magnetici hanno verso opposto, il loro riallineamento libera energia. Quindi, l'energia magnetica totale è minima se i momenti magnetici di tutti gli atomi hanno lo stesso verso.

Fornendo energia termica a un materiale ferromagnetico si può influire sull'allineamento degli spin. Se il ferro puro viene riscaldato ad alta temperatura, l'energia termica predomina sulle interazioni ferromagnetiche, cosicché il verso di ogni momento magnetico cambia casualmente da un istante all'altro. Un'istantanea degli atomi di ferro mostrerebbe che in media vi sono tanti momenti magnetici rivolti verso l'alto, quanti verso il basso, verso sinistra e verso destra, in avanti e all'indietro. Il vettore somma di tutti i momenti magnetici, o magnetizzazione netta, è nullo. (Per essere precisi, vi è una probabilità elevatissima che la magnetizzazione netta sia estremamente piccola.) In questa fase il ferro è chiamato paramagnetico.

Quando la temperatura viene ridotta, le interazioni tra i momenti magnetici diventano predominanti e i vettori tendono ad allinearsi in uno stato di energia inferiore. Alla temperatura critica di 771 gradi Celsius la disposizione degli atomi cambia repentinamente e radicalmente e i momenti magnetici si allineano per lo più nello stesso verso. (La magnetizzazione appare assente in un comune campione di ferro a causa di un altro complicato processo che suddivide lo stato ordinato in dominî. In ogni dominio, tuttavia, i momenti magnetici hanno tutti lo stesso verso.) Alla temperatura critica il ferro subisce una transizione di fase da paramagnetico a ferromagnetico.

In altri materiali, invece, prevale un diverso tipo di ordine negli stati di bassa energia. Per esempio, atomi di cromo contigui tendono ad allineare i momenti magnetici in versi opposti; se il momento magnetico di un atomo è orientato verso l'alto, quello di un atomo adiacente è orientato verso il basso. Questo comportamento, opposto a quello del ferro, viene definito antiferromagnetico. Analogamente a un materiale ferromagnetico, anche il cromo ha una temperatura critica alla quale si trasforma da materiale paramagnetico (con disposizione casuale dei momenti magnetici) a materiale antiferromagnetico (con i momenti magnetici allineati in versi opposti).

È notevole il fatto che i vetri di spin mostrino proprietà sia ferromagnetiche sia antiferromagnetiche. In certi vetri di

Cristallo e vetro, pur apparentemente molto simili, sono fasi strutturali distinte della materia. Un cristallo è un solido, mentre un vetro è un fluido a scorrimento lento. Analogamente, nei cosiddetti «vetri di spin» si può avere un'orientazione permanente degli spin degli atomi, il che dà luogo a una nuova fase della materia; oppure l'orientazione degli atomi può variare in modo estremamente lento.

spin, un metallo ospite non magnetico viene «diluito» con un piccolo numero di atomi che hanno momento magnetico non nullo, ossia sono dotati di spin. Due atomi magnetici adiacenti in una «lega magnetica» così diluita possono presentare interazioni ferromagnetiche o antiferromagnetiche. Per esempio, se si mescolano alcune parti di ferro con 100 parti di rame gli atomi di ferro, che di solito interagiscono in modo ferromagnetico, possono interagire anche in modo antiferromagnetico. Sebbene le ragioni di un comportamento di questo tipo siano da ricondurre alle sottigliezze della teoria dei quanti, il fenomeno può essere descritto qualitativamente.

Ciascuno degli elettroni di conduzione che si muovono liberamente nel rame possiede uno spin che viene influenzato in modo alquanto strano da un atomo di ferro. A una certa distanza gli atomi di ferro fanno orientare lo spin degli elettroni di conduzione in modo che sia parallelo al loro stesso spin; a una distanza poco maggiore lo spin degli elettroni di conduzione è antiparallelo allo spin del ferro. Ancora più in là lo spin è parallelo, e così via. Quindi l'atomo di ferro si trova al centro di una successione di sfere concentriche di influenza decrescente in cui il suo effetto è alternativamente ferromagnetico e antiferromagnetico.

Poiché gli elettroni di conduzione del metallo ospite mediano le interazioni tra due atomi magnetici adiacenti, questi ultimi possono interagire in modo ferromagnetico o antiferromagnetico a seconda della distanza tra di essi. Di conseguenza, in un vetro di spin composto da molti atomi di un metallo dispersi in una matrice di un altro metallo, circa metà di tutte le coppie di atomi interagirà in modo ferromagnetico, mentre l'altra metà interagirà in modo antiferromagnetico. Pertanto in metà dei casi l'energia magnetica di una coppia di atomi di ferro diminuisce se i loro spin sono paralleli, mentre negli altri casi l'energia diminuisce se gli spin sono antiparalleli.

Il risultato di questo duplice comportamento è che un atomo con un dato spin può non essere in grado di orientarsi in modo tale da interagire appropriatamente con tutti gli altri atomi del vetro di spin. Immaginiamo tre atomi di ferro distribuiti casualmente in un reticolo di rame. Il primo atomo interagisce in modo antiferromagnetico con il secondo, mentre le interazioni tra il primo e il terzo e tra il secondo e il terzo sono ferromagnetiche. Non esiste alcuna possibilità di realizzare contemporaneamente tutte le interazioni. Per esempio, se lo spin del primo atomo è orientato verso l'alto, lo spin del secondo deve essere orientato verso il basso. Il terzo, a sua volta, dovrebbe allineare il proprio spin nella stessa direzione sia del primo (spin verso l'alto) sia del secondo (spin verso il basso). È evidente che almeno una di queste interazioni è impossibile. Un sistema nel quale non possono manifestarsi contemporaneamente tutte le possibili interazioni si dice «frustrato».

Una diretta conseguenza della frustrazione è che un vetro di spin può avere molti stati di bassa energia. Per esempio, se il primo, il secondo e il terzo atomo hanno spin rispettivamente su, giù, su o su, giù, giù, essi si trovano nel loro stato energetico più basso, perché il numero di violazioni è il minimo possibile.

Le conseguenze della frustrazione si spingono ben oltre la fisica dei vetri di spin fino a problemi complessi in molte altre discipline. Nel caso dei vetri di spin, il fatto che non vi sia un unico stato di bassa energia è correlato alla domanda: il vetro di spin è una nuova fase della materia o è solo una sostanza paramagnetica estremamente viscosa? La transizione da liquido a cristallo o da comportamento paramagnetico a comportamento ferromagnetico al diminuire della temperatura è una vera transizione di fase: gli stati che ne risultano conservano un ordinamento distinto fino a quando la temperatura viene mantenuta entro certi valori. D'altra parte il vetro comune, sebbene appaia come una fase a sé stante, è fondamentalmente un liquido: esso fluisce, ma con una velocità tanto bassa da sembrare un solido.

Il vetro di spin potrebbe essere una fase distinta della materia, il cui ordinamento magnetico, o allineamento degli spin, si conserva fino a quando la temperatura viene mantenuta bassa. D'altra parte questi vetri potrebbero essere materiali paramagnetici le cui proprietà dinamiche sono rallentate al punto che essi sembrano costituire una fase statica. Se si osservasse un cambiamento di orientazione negli spin di un vetro di spin mantenuto a bassa temperatura si potrebbe concludere che esso è semplicemente un materiale paramagnetico. Per tale verifica definitiva, tuttavia, il vetro di spin dovrebbe venire osservato probabilmente per un periodo più lungo dell'età dell'universo.

Una cosa che si può fare in laboratorio è cercare di osservare una transizione di fase, ossia un'improvvisa variazione delle proprietà magnetiche e termodinamiche del vetro di spin in corrispondenza di una temperatura critica. Purtroppo le misurazioni di differenti proprietà hanno fornito risposte discordanti.

Nel 1970 Vincent D. Cannella, John A. Mydosh e Joseph I. Budnick della Fordham University studiarono le proprietà magnetiche di leghe oro-ferro. Tra l'altro essi misurarono la suscettività magnetica delle leghe, cioè la variazione di magnetizzazione al variare (per piccoli incrementi) di un campo magnetico esterno. Al pari di molti altri ricercatori che nel frattempo hanno ripetuto l'esperimento, essi osservarono una brusca variazione di suscettività a una temperatura critica al tendere a zero del campo magnetico. Questa improvvisa variazione indicava una transizione di fase.

Nei materiali paramagnetici, ferromagnetici e antiferromagnetici gli atomi si comportano come minuscole barrette magnetiche; le frecce indicano il «polo nord» di un atomo (cioè l'orientazione del suo spin). Gli atomi di un materiale paramagnetico (*a sinistra*) hanno un'orientazione degli spin casuale e variabile nel tempo. In un materiale ferromagnetico (*al centro*) gli spin sono tutti allineati nello stesso verso, mentre in un materiale antiferromagnetico (*a destra*) gli atomi adiacenti hanno spin allineati in versi opposti. Un vetro di spin è forse simile a un materiale paramagnetico «congelato», con un'orientazione degli spin casuale e permanente.

Altri esperimenti hanno condotto alla conclusione opposta. Ci si potrebbe aspettare che una transizione di fase sia contraddistinta da una brusca variazione di una grandezza termodinamica quale il calore specifico (la quantità di calore necessaria per far aumentare di un certo valore la temperatura di una sostanza), ma nessuna variazione del genere è stata osservata. Anzi, il calore specifico di molti tipi diversi di vetri di spin presenta un'ampia curva continua con un massimo a una temperatura che in genere è superiore di circa il 20 per cento alla temperatura di transizione indicata dalle misurazioni della suscettività.

Inoltre, certi dati fanno pensare che le misurazioni di proprietà magnetiche e termiche siano eseguite su tempi troppo brevi da permettere al vetro di spin di reagire completamente al cambiamento delle condizioni. Non è quindi ancora chiaro se sia stata effettivamente osservata in laboratorio una transizione di fase nei vetri di spin.

Mentre i ricercatori di laboratorio lamentano i risultati contraddittori, i teorici sono ostacolati dalla scarsa conoscenza della termodinamica dei sistemi disordinati. Nell'ultimo decennio il grosso della ricerca teorica è stato eseguito, per motivi di semplicità e di trattabilità, su modelli di vetri di spin chiaramente non realistici. Nonostante ciò, il tentativo di capire tali semplici modelli ha fornito notevoli risultati, rivelando molte caratteristiche sorprendenti.

Sam F. Edwards dell'Università di Cambridge e Philip W. Anderson della Princeton University hanno introdotto nel 1975 un modello «a breve raggio d'azione» dei vetri di spin che è stato molto studiato. Nel loro modello, gli spin erano posizionati ai vertici di un reticolo cubico. Uno spin interagiva solo con i suoi vicini immediati, ma una data interazione aveva la stessa probabilità di essere ferromagnetica o antiferromagnetica. Edwards e Anderson assegnarono a caso valori numerici alle interazioni. (I numeri indicavano l'intensità e il tipo di interazione ed erano positivi per le interazioni ferromagnetiche, negativi per quelle antiferromagnetiche.)

Supponiamo ora di tracciare una curva chiusa nel reticolo, cominciando con un'interazione tra uno spin e il suo vicino, spostandoci verso un'interazione adiacente e continuando così fino a ritornare all'atomo di partenza. Se registriamo i valori delle interazioni lungo la curva chiusa e li moltiplichiamo tra loro, si ha frustrazione ogni volta che il prodotto è negativo. Dal momento che un normale reticolo contiene molte curve chiuse frustrate, è estremamente difficile determinare, per tutti i siti, i valori degli spin che rendono minima l'energia totale del sistema. In realtà il sistema può avere molti stati di bassa energia, anche del tutto privi di correlazioni reciproche. Fu ben presto chiaro che i metodi matematici esistenti non erano in grado di rivelare il comportamento a bassa energia neppure di questo modello relativamente semplice.

Subito dopo le ricerche di Edwards e Anderson, David Sherrington dell'Imperial College di Londra e Scott Kirkpatrick del Thomas J. Watson Research Center dell'IBM proposero un modello «a raggio d'azione infinito». Essi ipotizzarono che in media ogni spin interagisse con la stessa intensità con ogni altro spin del sistema; la loro speranza era che l'ipotesi, pur essendo molto meno realistica, potesse condurre a un modello facilmente risolvibile. Il modello presentava una vera transizione di fase, da una fase paramagnetica a una di vetro di spin nella quale gli spin erano congelati in una disposizione permanente casuale. Comunque, i tentativi di ricavare le proprietà della fase di vetro di spin risultarono molto più difficili del previsto.

Nel 1979 Giorgio Parisi dell'Università di Roma ricavò una soluzione del modello di Sherrington-Kirkpatrick. Tale soluzione non è stata ancora dimostrata definitivamente, anche se vi sono numerosi dati che la sostengono. Tuttavia, all'epoca della proposta, essa appariva tanto misteriosa e diversa dalle precedenti descrizioni di fasi della materia che passarono quattro anni prima che le sue proprietà potessero essere comprese in senso fisico.

La soluzione di Parisi indica che, al di sotto della temperatura critica, un vetro di spin di Sherrington-Kirkpatrick può essere congelato in una fase fra le molte possibili, ciascuna delle quali rappresenta uno stato di bassa energia. La correlazione tra questi stati non è però una semplice trasformazione di simmetria, quale l'inversione di tutti gli spin. Per passare da uno stato energetico di bassa energia a un altro è necessario invertire una frazione significativa degli spin. Nel 1977 le ricerche di Anderson, David J. Thouless dell'Università di Birmingham e Richard G. Palmer di Princeton rafforzarono l'ipotesi che possano esistere molte soluzioni non banali in un modello di vetro di spin di grande scala. Questo aspetto sorprendente del modello di Sherrington-Kirkpatrick ha importanti conseguenze in problemi di scienza dei calcolatori, di biologia e di altre discipline, come vedremo ora.

In un vetro di spin si sviluppa «frustrazione» quando vi è competizione fra le interazioni magnetiche. Nella configurazione mostrata, le linee verdi e rosse rappresentano interazioni che allineano gli spin rispettivamente nello stesso verso e nel verso opposto. Le frecce rappresentano il verso degli spin in una disposizione che soddisfa tutte le interazioni tranne quella fra *B* e *C*. In questo caso non esiste una disposizione degli spin in grado di soddisfare tutte le interazioni.

Marc Mézard, Nicolas Sourlas e Gérard Toulouse dell'École Normale Supérieure di Parigi, Miguel A. Virasoro dell'Istituto Marconi di Roma e Parisi dimostrarono nel 1984 che questi stati di bassa energia presentano comunque alcune correlazioni. In particolare, gli stati si possono ordinare in modo gerarchico, come in un albero genealogico o in un dendrogramma filogenetico, sulla base delle distanze reciproche. Per determinare la distanza tra gli stati, si considera per ogni stato una mappa che fornisce l'orientazione di ciascuno spin. Le mappe sono sovrapposte a coppie per poter confrontare un dato spin di uno stato con lo spin dell'altro stato che si trova nella posizione corrispondente. Si sommano poi le differenze tra gli spin dei due stati per ricavare la distanza. Gli stati di bassa energia tendono a raggrupparsi insieme nell'albero genealogico: la minima distanza è quella tra «fratelli», seguono i «primi cugini», e così via. La comparsa di una simile struttura gerarchica è molto insolita nei sistemi fisici; essa ricorda piuttosto sistemi biologici e altri sistemi sottoposti a evoluzione e potrebbe essere un aspetto comune a molti sistemi disordinati.

Tuttavia i recenti studi dei più realistici modelli a breve raggio d'azione, compiuti dallo scomparso William L. McMillan dell'Università dell'Illinois a Urbana-Champaign, da Daniel S. Fisher e David A. Huse degli AT&T Bell Laboratories e da Alan J. Bray e Michael A. Moore dell'Università di Manchester,

L'«albero genealogico» dei vetri di spin descrive i rapporti tra gli stati di bassa energia della soluzione proposta da Giorgio Parisi dell'Università di Roma. Sovrapponendo due degli stati e contando gli spin orientati in versi opposti si ottiene la distanza tra gli stati. Le frecce in colore indicano che la distanza tra gli stati *A* e *D* è tre. Il primo livello dell'albero collega stati separati da una distanza di due, il secondo livello stati separati da una distanza di tre.

sembrano indicare che la struttura gerarchica è un aspetto patologico del modello di Sherrington-Kirkpatrick a raggio d'azione infinito. Anche se la validità di queste nuove teorie è ancora in discussione, esse riproducono molte delle proprietà dei vetri di spin osservate in laboratorio. Queste teorie prevedono che una fase di vetro di spin a bassa temperatura sia formata da due soli stati di minima energia; una visione quindi del tutto differente da quella del modello di Sherrington-Kirkpatrick. La spiegazione di come i modelli a breve raggio d'azione e a raggio d'azione infinito siano correlati tra loro e ai vetri di spin reali rimane una difficile sfida per la comunità dei fisici teorici.

Nonostante la relativa ignoranza sulla natura dei vetri di spin, nell'ultimo decennio sono stati compiuti tentativi audaci, entusiasmanti e forse perfino avventati di applicazione delle attuali conoscenze ad alcuni fondamentali problemi di scienza dei calcolatori, neurologia e biologia. Quali potrebbero essere le caratteristiche in comune tra i vetri di spin e questi problemi? Molti di essi non si possono ridurre a problemi a poche variabili; anzi, come i vetri di spin, essi comprendono un gran numero di variabili che interagiscono in modo non uniforme. Nella maggior parte dei casi, come accade con la frustrazione nei vetri di spin, i vincoli del problema non possono essere soddisfatti tutti contemporaneamente; perciò esistono spesso molte soluzioni possibili, apparentemente non correlate.

I vetri di spin hanno fornito uno schema matematico che possiede tutte queste interessanti proprietà e altre ancora. La teoria dei vetri di spin è stata perciò un ovvio punto di partenza per costruire modelli rudimentali di questi altri sistemi complessi. Essa ha permesso di simulare talune proprietà dei sistemi che finora era difficile o impossibile descrivere con un modello.

Alcune delle prime applicazioni della matematica dei vetri di spin hanno assunto la forma di algoritmi per calcolatore destinati a risolvere problemi di ottimizzazione combinatoria. Un esempio ben noto di tali problemi è quello del commesso viaggiatore. Qual è la distanza più breve che un commesso può percorrere per visitare un certo numero di città e ritornare alla città di partenza? La risposta può sembrare elementare: si calcola la distanza tra ogni coppia di città, si sommano le distanze per tutte le possibili combinazioni di coppie che completano il viaggio e si sceglie la combinazione più breve. Anche se questa strategia può andar bene nel caso che le città siano poche, al di là di un certo numero iniziano a sorgere problemi; all'aumentare del numero di città, il numero di possibilità diventa enorme.

Per esempio, se il numero di città è cinque, un calcolatore può calcolare facilmente le 12 diverse possibilità. Con 10 città, è ancora possibile calcolare le 181 440 possibilità. Tuttavia, per sole 25 città, il numero di possibili viaggi è tanto grande che un calcolatore capace di valutare un milione di possibilità al secondo impiegherebbe 9,8 miliardi di anni, quasi due terzi dell'età dell'universo, per esaminarle tutte.

Di solito problemi di ottimizzazione combinatoria come questo comprendono un gran numero di variabili e di vincoli, un numero enorme di possibili combinazioni e una funzione «costo» che descrive tutti i possibili valori della grandezza da ottimizzare. (Nel problema del commesso viaggiatore la funzione costo è la lunghezza del viaggio in funzione di ogni possibile percorso.) Posti in termini semplici, i problemi di ottimizzazione combinatoria si riducono alla domanda: qual è la soluzione di minor costo?

Per molti di questi problemi, anche quelli che hanno un gran numero di variabili, algoritmi geniali riescono a trovare in un tempo relativamente breve la soluzione di costo minimo, o «globalmente» ottimale. Per certi tipi di problemi di ottimizzazione combinatoria, tuttavia, si ritiene fermamente che sia impossibile escogitare un algoritmo che trovi la soluzione migliore in un ragionevole lasso di tempo per ogni caso del problema. Questi problemi «difficili» hanno spesso applicazioni pratiche nella logica, nella robotica, nella teoria del linguaggio e nell'immagazzinamento e recupero di dati. Non deve quindi sorprendere il fatto che uno di questi problemi difficili consista nel trovare la disposizione di spin di minima energia che soddisfi il modello tridimensionale dei vetri di spin di Edwards-Anderson.

Una strategia alternativa per problemi difficili di ottimizzazione combinatoria è quella di cercare soluzioni «localmente» ottimali. Queste soluzioni di costo relativamente basso non possono essere migliorate con nessun piccolo riassestamento dei valori assegnati, come l'inversione di alcuni spin in un vetro di spin o la variazione dell'ordine in cui vengono visitate alcune città nel problema del commesso viaggiatore. Anche se può esservi una certa riluttanza ad accettare una «buona» soluzione localmente ottimale, l'alternativa è forse quella di attendere fino al prossimo secolo un calcolatore capace di ricavare «di forza» la soluzione migliore.

Per coloro, quindi, che non sono disposti ad attendere tanto a lungo, Kirkpatrick, Charles D. Gelatt, Jr., e Mario Vecchi del Thomas J. Watson Research Center dell'IBM hanno inventato uno strumento di calcolo. Essi hanno progettato un algoritmo per calcolatore, chiamato «ricottura simulata», che trova soluzioni localmente ottimali in tempi ragionevolmente brevi. Questo algoritmo si basa su metodi fisici per ottenere uno stato di bassa energia in un vetro di spin.

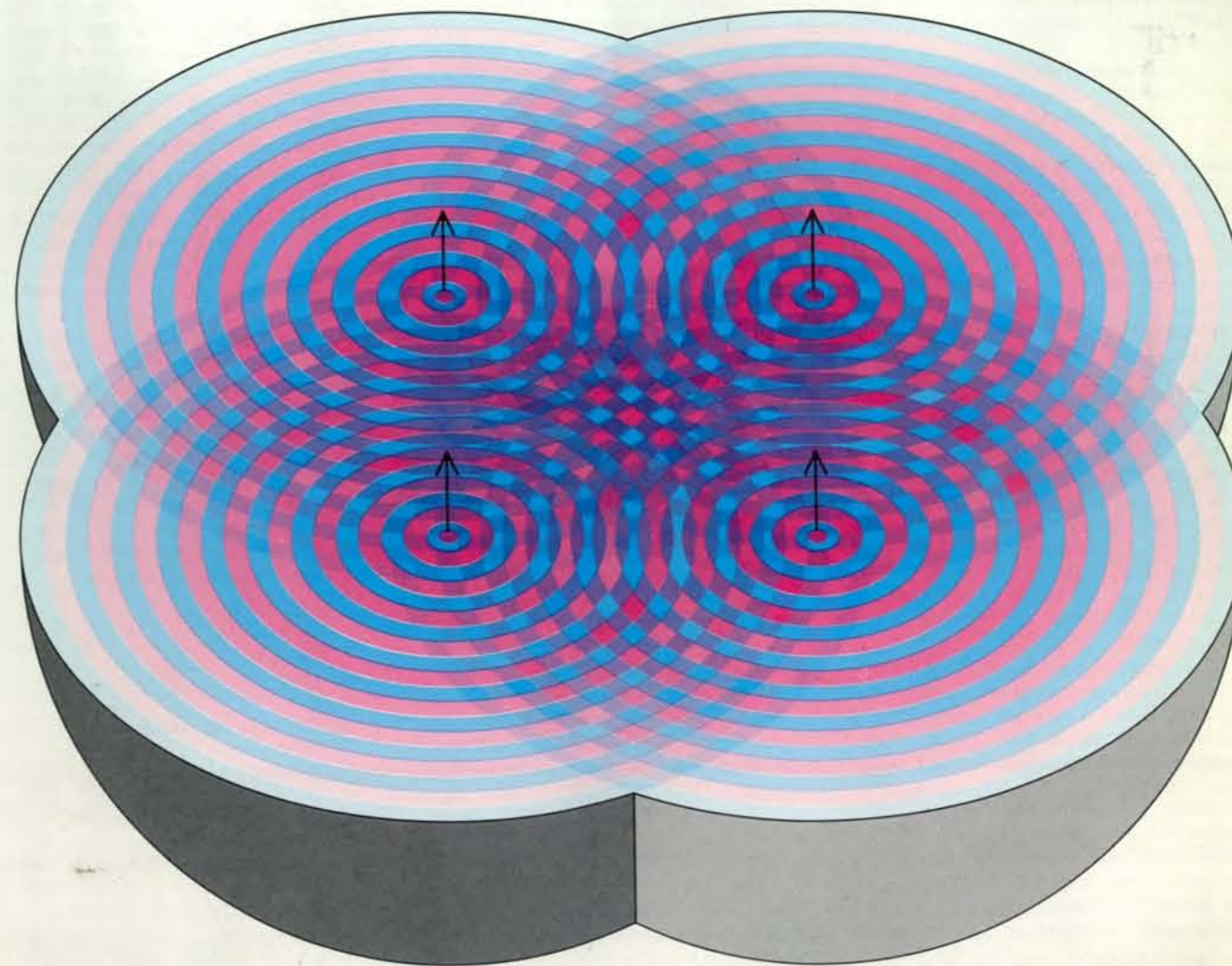
L'energia di un particolare vetro di spin può essere descritta in funzione dello stato del sistema, ossia di una particolare disposizione degli spin. Se ogni spin può essere orientato in una direzione su due (per esempio, in alto o in basso) il numero totale di stati distinti è due elevato al numero di spin del sistema. Il sistema definisce se ogni data coppia di spin interagisce in modo ferromagnetico o antiferromagnetico e l'energia di un particolare stato dipende da quanti spin di tale stato soddisfano le interazioni. Se queste sono tutte soddisfatte lo stato ha energia minima; in caso contrario, ha un valore di energia superiore.

L'energia come funzione di tutti gli stati descrive una superficie in uno spazio le cui dimensioni sono uguali al numero degli spin. Invece di cercare di visualizzare tale geometria, si immagini la funzione come una catena di montagne: la quota di ogni punto rappresenta l'energia di ciascuno stato. Supponiamo che il sistema si trovi attualmente in un particolare stato di alta energia: possiamo immaginarlo come un atleta su una bicicletta da montagna in cima a una delle vette. L'obiettivo dell'atleta è la ricerca della valle più profonda - lo stato di minima energia - senza avere a disposizione una mappa.

L'atleta scende costeggiando il fianco della montagna fino a raggiungere una valle vicina. Ritenendo improbabile di poter trovare la valle più profonda al primo tentativo, esce pedalando dalla valle, raggiunge un passo montano e ridiscende in una valle ancora più profonda. Dopo aver pedalato per molte valli e molti passi, l'atleta crede di aver trovato la valle più profonda, anche se non può averne la certezza, poiché molte valli sono rimaste inesplorate.

La ricerca di uno stato di bassa energia di un vetro di spin richiede un processo di riscaldamento e raffreddamento - ossia di ricottura - analogo alle salite e alle discese dell'atleta. Se la temperatura è estremamente bassa, il sistema rimane in una valle anche poco profonda per un periodo molto lungo. Si potrebbe dire che all'aumentare della temperatura il sistema disponga di più energia per «compiere esplorazioni». L'orientazione degli spin può invertirsi facilmente e quindi esso ha una più elevata probabilità di uscire dalle valli meno profonde e può cercare molte altre configurazioni possibili di spin, alcune delle quali potrebbero avere energia inferiore a quella dello stato di partenza.

Un semplice algoritmo per la ricerca di uno stato di energia relativamente bassa di un vetro di spin consiste quindi nel simulare un'alta temperatura (in modo che il sistema, in linea di principio, possa cercare uno stato qualsiasi) e nel raffreddare lentamente il sistema in modo che possa assestarsi in uno stato di energia inferiore. Se esso si blocca temporaneamente in una valle ad alta quota ha ancora una buona probabilità di sfuggire dal passo più vicino andando alla ricerca di una valle più profonda (di minore energia). Dopo molti cicli di riscaldamento e di raffreddamento, l'algoritmo ha una elevata probabilità di fornire una buona soluzione - cioè uno stato di bassa energia - anche se la probabilità di trovare accidentalmente la migliore soluzione in uno spazio così grande è estremamente piccola.

In molti problemi di ottimizzazione combinatoria la funzione costo rassomiglia a un paesaggio tormentato in uno spazio di stati, proprio come la funzione energia nel caso di un vetro di spin. Gli algoritmi di ricottura simulata affrontano i problemi di ottimizzazione combinatoria come se l'obiettivo fosse quello di trovare una soluzione di bassa energia per un vetro di spin. La funzione costo

Un vetro di spin può avere origine dall'interazione di atomi con elettroni di conduzione. Le frecce indicano il verso dello spin di un atomo. Ogni atomo si trova al centro di una serie di sfere concentriche di influenza decrescente (*in colore*), nelle quali gli spin degli elettroni che circondano l'atomo hanno alternativamente allineamento antiparallelo (*in rosso*) e parallelo (*in blu*) allo spin dell'atomo. Gli elettroni mediano le interazioni fra gli atomi, i cui spin possono invertirsi per effetto degli atomi e degli elettroni circostanti.

è l'analogo dell'energia come funzione di stato. Anche se in questi problemi la temperatura non ha alcun significato fisico, essa può essere aumentata o diminuita come una temperatura reale, consentendo così al sistema di esplorare regioni dello spazio degli stati alla ricerca di una buona soluzione. In tal modo gli algoritmi di ricottura simulata possono trovare con relativa velocità soluzioni a basso costo per un gran numero di problemi di ottimizzazione combinatoria.

All'inizio degli anni ottanta, John J. Hopfield degli AT&T Bell Laboratories e del California Institute of Technology propose un'altra importante applicazione della matematica dei vetri di spin. Egli comprese che un sistema simile ai vetri di spin poteva eseguire calcoli e immagazzinare informazioni, purché fosse fornito delle opportune regole dinamiche. Il sistema era particolarmente interessante perché imitava l'architettura del cervello più strettamente dei normali calcolatori digitali (si veda l'articolo *Circuiti elettronici basati su modelli biologici* di David W. Tank e John J. Hopfield in «Le Scienze» n. 234, febbraio 1988).

Il modello di Hopfield è costituito da semplici «neuroni», ciascuno dei quali può assumere uno stato su due possibili: acceso (durante la scarica) e spento (in quiete). Il fatto che un neurone rimanga nel suo stato attuale o cambi stato dipende dagli stati di tutti i neuroni a esso collegati. La natura del calcolo da effettuare determina lo schema dei collegamenti tra i neuroni. Al contrario dei veri neuroni, quelli del modello di Hopfield interagiscono simmetricamente: l'effetto di un neurone su un altro è analogo all'effetto del secondo sul primo.

Sotto alcuni aspetti questo sistema rassomiglia a un vetro di spin: è un insieme di variabili - ciascuna delle quali può assumere due valori - che interagiscono in modo complesso e non uniforme. Vi sono però differenze fondamentali nel senso che la distribuzione di probabilità delle interazioni del sistema neuronico è scelta a priori e le interazioni possono evolvere nel tempo, il che consente l'«apprendimento».

Per questo sistema si può definire una funzione energia dipendente dallo stato allo stesso modo che per i vetri di spin. Come in precedenza, il risultato è un tormentato paesaggio di montagna nello spazio degli stati. Le valli di energia potrebbero corrispondere a ricordi da richiamare, a forme da riconoscere o ad altri tipi di processi mentali. I collegamenti tra i neuroni del sistema stabiliscono il numero, la posizione e il significato delle valli.

Uno stimolo esterno determina la posizione iniziale del sistema nello spazio degli stati, ossia stabilisce quali neuroni sono in attività e quali in quiete. Un bacino di attrazione circonda il punto più basso di una valle: la «soluzione». Se il sistema si trova in un punto qualsiasi del bacino, esso evolve verso la soluzione associata. Pertanto, i segnali provenienti dall'esterno danno avvio alla scelta di una data soluzione, richiamando, per esempio, un dato ricordo.

Questo tipo di sistema differisce notevolmente dagli algoritmi lineari, sequenziali, progettati per i calcolatori digitali. Il tipo di calcolo qui impiegato, come quello che controlla il sistema nervoso di un animale, è collettivo, vale a dire tutte le parti del sistema interagiscono simultaneamente l'una con l'altra e molte di esse cambiano di stato durante lo svolgimento del calcolo.

Un'altra connessione tra la matematica dei vetri di spin e la biologia riguarda il campo dell'evoluzione biologica. Una domanda fondamentale relativa agli inizi dell'evoluzione è la seguente: come può un «brodo» di piccole molecole come amminoacidi o acidi nucleici dare origine a macromolecole, quali le proteine e il DNA, altamente organizzate e in grado di portare informazione? I processi chimici che hanno determinato questa evoluzione sono sconosciuti, tuttavia è possibile costruire un modello matematico dell'evoluzione molecolare in cui si verifica un'interessante transizione da un piccolo numero di «informazioni» a numero superiore.

Per esempio, le informazioni biologiche contenute in una sequenza di monomeri (i «mattoni» delle macromolecole o polimeri) hanno valore se aiutano il polimero a sopravvivere. Prima che esistesse un complesso meccanismo cellulare atto a tradurre le sequenze di DNA in proteine funzionali, la probabilità di sopravvivenza di polimero era forse direttamente correlata alle proprietà chimiche della sequenza in sé: facilità di replicazione, caratteristiche di piegamento, probabilità di essere assorbita su una superficie contigua, tendenza all'autoaggrovigliamento, stabilità e così via.

All'inizio degli anni ottanta Daniel S. Rokhsar di Princeton, Anderson e io abbiamo preso in esame due monomeri, X e Y, che in linea di principio possono rappresentare qualsiasi coppia di molecole di interesse biologico. I due monomeri vengono mescolati in parti uguali. Al sistema viene imposta una successione di eventi che porta alla formazione di filamenti sempre più lunghi di monomeri al passare del tempo. La chiave per fornire un ricco contenuto di informazione sta in un processo caratterizzato sia da diversità sia da selezione. Diversità significa che questa successione di eventi può produrre un gran numero di polimeri differenti. Se, per esempio, un legame X-Y è molto più stabile di un legame X-X o di uno Y-Y, il processo di crescita produce sempre polimeri con la sequenza X-Y-X-Y-X-Y e così via, e dal processo è possibile ricavare ben poche informazioni. Selezione vuole dire invece che non tutti i polimeri hanno la stessa probabilità di formarsi o di sopravvivere. Se un sistema produce ogni possibile tipo di polimeri, di nuovo non si ottiene alcuna informazione. Il contenuto di informazione di questo sistema può essere definito semplicemente dal logaritmo del rapporto tra il numero di polimeri possibili e il numero di quelli esistenti.

Per creare sia diversità sia selezione i miei colleghi e io abbiamo studiato un modello abbastanza astratto nel quale l'interazione tra due monomeri di una catena polimerica ha una probabilità circa uguale di far aumentare o di far diminuire la probabilità di sopravvivenza della sequenza. Per una data catena, la probabilità totale di sopravvivenza per unità di tempo è la somma dei contributi di ogni coppia. Si può costruire un modello che spieghi statisticamente questi fattori; da un punto di vista matematico, esso risulta molto simile ai modelli dei vetri di spin.

Come la funzione di stato dei vetri di spin, anche la funzione probabilità di sopravvivenza presenta molti picchi e valli nello spazio degli stati di tutti i polimeri del modello. Il modello dimostra che non tutti i polimeri sono ugualmente probabili (cosa che pare realistica). Inoltre, via via che il numero dei polimeri «buoni» aumenta esponenzialmente con le dimensioni dei polimeri, viene mantenuta una diversità sufficiente a produrre informazione.

Qualunque risultato possa derivare da queste ricerche, esse hanno favorito un apprezzabile scambio di contributi tra discipline diverse quali fisica, matematica, scienza dei calcolatori, biologia, chimica, economia. Un numero sempre maggiore di ricercatori sta affrontando problemi fondamentali che esemplificano l'incertezza e il disordine dell'universo. Per ora abbiamo cominciato a comprendere i motivi per cui questi sistemi sono così difficili da trattare matematicamente e forse potremo arrivare a conoscerli meglio nei loro stessi termini. Finalmente ci siamo decisi a rimboccarci le maniche e a sporcarci le mani.

BIBLIOGRAFIA

MÉZARD M., PARISI G. e VIRASORO M. A., *Spin Glass Theory and Beyond*, World Scientific Publications, Singapore, 1986.

CHOWDHURY D., *Spin Glasses and Other Frustrated Systems*, World Scientific Publications, Singapore, 1986.

STEIN D. L. (a cura), *Complex Systems*, Addison-Wesley Publishing Company, in stampa.

KAUFFMAN S. A., *Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, in stampa.

Reazioni chimiche oscillanti

Un tempo ritenute contrarie alle leggi naturali, è stato invece possibile realizzarle di recente in laboratorio; sembra che possano chiarire analoghi comportamenti nei sistemi viventi

di Irving R. Epstein, Kenneth Kustin, Patrick De Kepper e Miklós Orbán

Ovunque in fisica, in astronomia e in biologia si trovano fenomeni oscillanti, o periodici. Essi vanno dal familiare movimento del pendolo e delle orbite dei pianeti ai complessi orologi biologici che governano il comportamento giornaliero e stagionale degli organismi viventi. Fino ad epoca recente, invece, i chimici credevano che le reazioni che avvengono nelle provette e nei bicchieri da laboratorio fossero singolarmente immuni da quell'andamento periodico che è così familiare in altri settori della scienza. Infatti la maggior parte dei chimici formatisi prima del 1950 era probabilmente convinta che, per una miscela di sostanze inorganiche semplici, il partecipare a una reazione che oscillava in modo visibile e periodico sarebbe stato una evidente violazione di una immutabile legge di natura. Anche oggi si suppone abitualmente che le reazioni chimiche siano unidirezionali: se due sostanze reagiscono per darne una terza, ci si aspetta che la reazione continuerà regolarmente fino all'esaurimento dei reagenti o al raggiungimento di uno stato di equilibrio. Normalmente non ci si aspetta che le concentrazioni dei prodotti intermedi della reazione raggiungano un certo livello, quindi scendano a un livello inferiore per risalire poi e di nuovo abbassarsi ripetutamente, finché si formano prodotti stabili resistenti a successivi cambiamenti.

Benché resoconti di tali reazioni apparissero sporadicamente nella letteratura chimica della fine dell'Ottocento e degli inizi del Novecento, essi furono respinti dalla larga maggioranza dei chimici come fenomeni non riproducibili, che potevano essere attribuiti probabilmente a processi estranei, come la corrosione, o la formazione di pellicole, che avvenivano nel corso della reazione. Negli ultimi 25 anni, tuttavia, lo studio delle reazioni chimiche oscillanti ha finalmente conquistato rispettabilità e ora costituisce uno dei campi della chimica in più rapido sviluppo. Lo studio promette di penetrare nella dinamica chimica e nei meccanismi di catalisi e forse nei fenomeni periodici ancora oscuri che si osservano in biologia e in geologia. Mentre il primitivo piccolo gruppo di reazioni chimiche oscillanti fu trovato per caso, è ora possibile descrivere le condizioni che contribuiscono all'oscillazione e preparare una ricerca sistematica di nuovi sistemi oscillanti. Si tratta di una ricerca che ha già dato i suoi frutti.

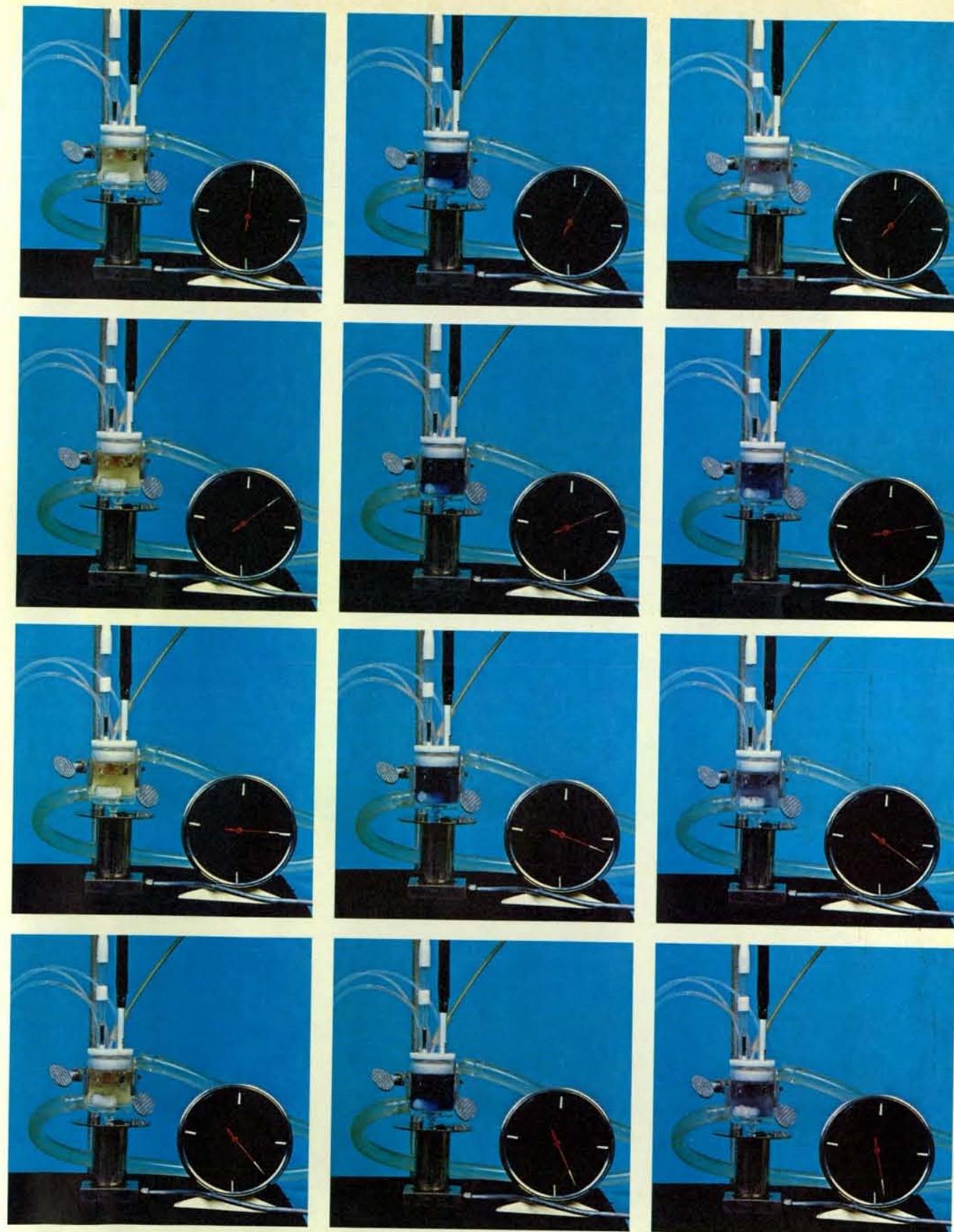
La riluttanza dei chimici ad accettare la realtà delle reazioni oscillanti può essere fatta risalire principalmente al secondo principio della termodinamica. Nella sua formulazione più nota, dovuta al fisico tedesco Rudolf Clausius nel XIX secolo, il secondo principio stabilisce che l'entropia, ossia la casualità, dell'universo tende ad aumentare. Applicato alle reazioni chimiche, il principio richiede che un sistema chimico, in assenza di apporti esterni di materia o di energia, deve continuamente tendere a uno stato finale di equilibrio. Cioè, se *A* va verso *B*, deve essere così per tutta la durata della reazione senza deviazioni a ritroso verso *A* lungo il percorso. Si pensava che le reazioni in cui questa regola sembra essere violata dovessero derivare o da uno scarso controllo delle condizioni sperimentali, o da qualche artificio voluto, dal momento che esse costituirebbero una sorta di macchina chimica a moto perpetuo.

Si può pertanto immaginare l'indifferenza con cui fu accolta la relazione su una reazione oscillante pubblicata nel 1921 da William C. Bray dell'Università della California a Berkeley. Studiando il ruolo dello iodato, ione dello iodio ossigenato (IO_3^-), come catalizzatore della decomposizione dell'acqua ossigenata, o perossido di idrogeno, in acqua e ossigeno, Bray osservò che, in determinate condizioni, la velocità di produzione dell'ossigeno e la concentrazione dello iodio nella soluzione cambiavano periodicamente. All'osservazione di Bray fu prestata scarsa attenzione nel corso dei successivi cinquant'anni. I pochi articoli che apparvero sulla reazione di Bray tentavano in gran parte di spiegare le oscillazioni come artefatti dovuti a polvere o a impurità più che di esaminare ulteriormente il fenomeno.

Come vedremo, lo iodio e gli ioni ioduro (I^-) e iodato svolgono un ruolo importante in molti sistemi oscillanti messi a punto di recente. L'era moderna per quanto riguarda lo studio di reazioni chimiche oscillanti può essere fatta partire da una scoperta accidentale dovuta nel 1958 al chimico sovietico B. P. Belousov. Egli osservò che, se si sciolgono in acqua acido citrico e acido solforico con bromato di potassio e un sale di cerio, la soluzione varia periodicamente da incolore a color giallo chiaro. Col 1958, dunque, alcuni chimico-fisici, almeno, furono disposti a considerare seriamente la notizia, sia perché la reazione di Belousov era facilmente ripetibile, sia perché i concetti di termodinamica del XIX secolo erano stati estesi dopo la seconda guerra mondiale.

Il primo in questo impegno fu Ilya Prigogine della Université libre di Bruxelles, il quale riconobbe che la termodinamica classica di Clausius richiedeva non solo che i sistemi fossero isolati dall'ambiente, ma anche che fossero vicini al loro stato di equilibrio. Per sistemi lontani dall'equilibrio, sia perché una reazione è solo ai primi stadi, sia perché il sistema è «aperto», e pertanto soggetto a un flusso di energia dall'esterno, Prigogine e i suoi collaboratori svilupparono il concetto di una termodinamica irreversibile. Per questo lavoro, nel 1977, Prigogine ha ricevuto il premio Nobel per la chimica.

In sistemi lontani dall'equilibrio, possono presentarsi molti nuovi fenomeni, descritti come strutture dissipative. Tali strutture comprendono oscillazioni periodiche nelle concentrazioni di specie intermedie in una reazione chimica; invece, non sono soggetti a oscillazione né i reagenti iniziali, né i prodotti finali. I viventi costituiscono gli esempi più interessanti e vari di sistemi oscillanti aperti, lontani dall'equilibrio. Essi vengono mantenuti in uno stato di non equilibrio dall'apporto di reagenti (sostanze nutritive) provenienti dall'esterno e dalla produzione di sostanze di rifiuto. Se l'uno o l'altro di questi flussi si arresta, l'organismo e le sue oscillazioni si estinguono.

Nel 1958, le implicazioni della termodinamica irreversibile di Prigogine furono

Un sistema chimico oscillante è fotografato nel laboratorio degli autori alla Brandeis University. La reazione procede in un recipiente di vetro nel quale vengono pompate tre soluzioni a una velocità costante: iodato di potassio, acido perclorico e acqua ossigenata, e acido malonico con solfato manganoso monoidrato. L'amido serve da indicatore, poiché in presenza di iodio forma un complesso blu scuro. L'intervallo fra una fotografia e la successiva, partendo dall'alto a sinistra e procedendo da sinistra verso destra, è poco meno di tre secondi (come è indicato dalla lancetta dei secondi sull'orologio). Questo particolare sistema è stato progettato da due insegnanti di scuola superiore in California, Thomas S. Briggs e Warren C. Rauscher. La reazione è stata modificata per farla avvenire in condizioni di flusso continuo.

capite soltanto da pochi chimici. Inoltre, la scoperta di Belousov passò quasi inosservata poiché era stata pubblicata in una raccolta sovietica poco nota di estratti tecnici sul ruolo delle radiazioni in medicina. La reazione oscillante di Belousov presentava importanti vantaggi rispetto al primitivo sistema di Bray. Essa avveniva a temperatura ambiente e non generava prodotti nocivi. Inoltre, le oscillazioni erano chiaramente visibili quando il cerio passava continuamente da uno stato ossidato (cerico), di colore giallo, a uno stato meno ossidato (ceroso), che è incolore.

Pochi anni dopo la sua pubblicazione, la reazione di Belousov attirò l'attenzione di A. M. Zhabotinsky dell'Istituto di fisica biologica nei pressi di Mosca. Egli apportò modifiche ai reagenti di Belousov (sostituendo, ad esempio, il sale di cerio con un sale di ferro che mostra un viraggio di colore dal rosso al blu, più evidente) e iniziò uno studio sistematico della reazione, che ora è comunemente chiamata reazione BZ (da Belousov-Zhabotinsky). All'inizio degli anni sessanta Zhabotinsky pubblicò un'ampia messe di risultati sperimentali, comprendenti la scoperta che, se si lascia indisturbato uno strato sottile di soluzione rossa BZ, inizialmente uniforme, compaiono ben presto dei punti blu che si aggregano fino ad assumere una sorprendente disposizione di strutture ad anello.

La conoscenza della reazione BZ (e della termodinamica irreversibile di Prigogine) si diffuse rapidamente sul finire degli anni sessanta e all'inizio degli anni settanta. Benché i chimici di tutto il mondo fossero affascinati dal fenomeno BZ, la grande maggioranza lo considerava ancora una curiosità, più utile per sorprendere i partecipanti a convegni, che come oggetto di studi seri. In breve tempo furono sviluppate molte varianti della reazione BZ, sostituendo uno o più componenti con specie chimiche molto affini per ottenere una reazione oscillante più lenta, più veloce, o diversamente colorata. Tra le varianti, quella che visivamente colpì di più fu un ibrido dei sistemi BZ e Bray, messo a punto nel 1973 da due insegnanti di una scuola superiore di San Francisco, Thomas S. Briggs e Warren C. Rauscher. Il loro sistema, che comprende acqua ossigenata, iodato di potassio, acido perclorico, acido malonico, solfato di manganese e amido, varia da incolore, al colore oro e al blu, per poi ricominciare da capo.

Benché la reazione BZ e le altre da questa derivate fossero diventate ben note verso la fine degli anni sessanta, grazie a dimostrazioni effettuate in ambiente scientifico, la natura del fenomeno oscillatorio restava un mistero. Un chimico capisce una reazione solo quando può scriverne il meccanismo, cioè una serie di reazioni, chiamate stadi elementari, ciascuna delle quali descrive un reale incontro tra molecole. La familiare equazione «stechiometrica», che i chimici sono soliti scrivere per descrivere una reazione, mostra solo il risultato netto e non il meccanismo con cui la reazione procede effettivamente su scala molecolare. Ad esempio, quando l'acetilene (C_2H_2) si combina con l'ossigeno (O_2) in un cannello per saldatura per formare anidride carbonica (CO_2) e acqua (H_2O), la reazione globale può essere scritta come $2\,C_2H_2 + 5\,O_2 \rightarrow 4\,CO_2 + 2\,H_2O$.

È però molto improbabile che due molecole di acetilene e cinque molecole di ossigeno vengano a collidere simultaneamente e si scindano, dando origine a quattro molecole di anidride carbonica e a due di acqua. La reazione in realtà avviene mediante una sequenza di stadi elementari, ciascuno dei quali comporta una collisione tra due molecole, o la scissione di una molecola singola, spesso con formazione o distruzione di specie intermedie che non compaiono affatto nell'equazione della reazione netta. Chiarire un tale meccanismo rappresenta una enorme sfida che richiede ingegnosità, l'analisi di molti dati sperimentali e spesso un'estesa simulazione al calcolatore.

Agli inizi degli anni settanta, si assunsero il formidabile compito di costruire un meccanismo per la reazione oscillante BZ Richard M. Noyes dell'Università dell'Oregon, in collaborazione con Richard J. Field (che è ora all'Università del Montana), ed Endre Körös dell'Università Eötvös di Budapest. Entro il 1972 essi riuscirono a ideare uno schema composto di 18 stadi elementari, che comportava circa 20 specie chimiche diverse e che sembrava in grado di spiegare le oscillazioni. Due anni dopo, una simulazione dettagliata al calcolatore svolta da Noyes, Field e David Edelson dei Bell Laboratories confermò che quel meccanismo prevede le oscillazioni. Studi successivi hanno dimostrato che esso spiega bene anche lo sviluppo di strutture spaziali. Noyes ha paragonato il suo metodo per chiarire il meccanismo al «metodo di [Sherlock] Holmes: "Quando le altre possibilità vengono meno, ciò che rimane, *per quanto improbabile,* deve essere la verità"».

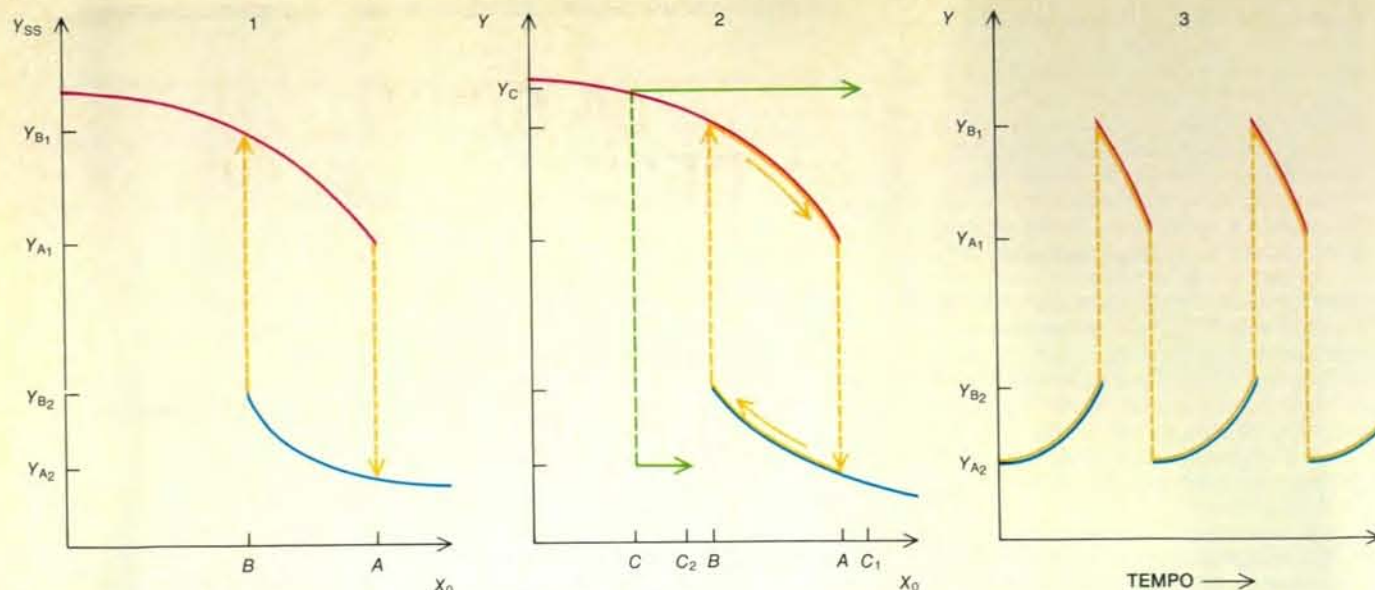
Con la soluzione del meccanismo BZ, le reazioni chimiche oscillanti divennero un'area importante e stimolante per coloro che erano interessati ai meccanismi di reazione e alla dinamica chimica. Furono avanzate varie ipotesi sulle condizioni richieste perché ci sia oscillazione in una reazione chimica e in alcuni ambienti nacque l'ottimistica speranza che la comprensione dell'oscillazione chimica potesse gettare nuova luce sul comportamento periodico degli organismi viventi. Nessuno, tuttavia, riuscì a definire una serie di condizioni necessarie e sufficienti per l'oscillazione chimica. Quando gli anni settanta giunsero al termine, il numero di reazioni chimiche diverse con andamento oscillante e allora note era terribilmente scarso; tutto era stato scoperto per caso; si conosceva una sola reazione e nessuno sapeva come scoprirne altre.

Anche se i sistemi biologici sono notoriamente una fonte fertile di preparati che mostrano oscillazione, essi tendono a sottrarsi al tipo di analisi meccanicistica che porta a una generalizzazione. Tali sistemi oscillanti possono essere ottenuti, per esempio, estraendo il contenuto chimico delle cellule di lievito e separandolo dalle pareti cellulari e da altre strutture estra-

Il reattore a flusso e agitazione continui ha facilitato la ricerca degli autori sulle reazioni chimiche oscillanti, fornendo le condizioni di non equilibrio necessarie perché abbiano luogo tali reazioni. L'illustrazione mostra una tipica disposizione sperimentale. I reagenti, contenuti in tre recipienti, vengono pompati in modo continuo alla base del recipiente di reazione. Nel primo sistema oscillante scoperto dagli autori, i materiali introdotti erano soluzioni acquose acide di un clorito, di uno iodato e di un arsenito. Le variazioni periodiche di colore sono controllate mediante la fotocellula. Le oscillazioni nella concentrazione di ioni particolari (in questo caso ioduri) sono rilevate da un elettrodo.

1	$2H^+ + Br^- + BrO_3^- \rightleftharpoons HOBr + HBrO_2$
2	$H^+ + HBrO_2 + Br^- \rightleftharpoons 2HOBr$
3	$HOBr + Br^- + H^+ \rightleftharpoons Br_2 + H_2O$
4	$CH_2(COOH)_2 \rightleftharpoons (OH)_2C{=}CHCOOH$
5	$Br_2 + (OH)_2C{=}CHCOOH \rightleftharpoons H^+ + Br^- + BrCH(COOH)_2$
6	$HBrO_2 + BrO_3^- + H^+ \rightleftharpoons 2BrO_2 + H_2O$
7	$BrO_2 + Ce^{3+} + H^+ \rightleftharpoons Ce^{4+} + HBrO_2$
8	$Ce^{4+} + BrO_2 + H_2O \rightleftharpoons BrO_3^- + 2H^+ + Ce^{3+}$
9	$2HBrO_2 \rightleftharpoons HOBr + BrO_3^- + H^+$
10	$Ce^{4+} + CH_2(COOH)_2 \rightleftharpoons CH(COOH)_2 + Ce^{3+} + H^+$
11	$CH(COOH)_2 + BrCH(COOH)_2 + H_2O \rightleftharpoons Br^- + CH_2(COOH)_2 + HOC(COOH)_2 + H^+$
12	$Ce^{4+} + BrCH(COOH)_2 + H_2O \rightleftharpoons Br^- + HOC(COOH)_2 + Ce^{3+} + 2H^+$
13	$2HOC(COOH)_2 \rightleftharpoons HOCH(COOH)_2 + O{=}CHCOOH + CO_2$
14	$Ce^{4+} + HOCH(COOH)_2 \rightleftharpoons HOC(COOH)_2 + Ce^{3+} + H^+$
15	$Ce^{4+} + O{=}CHCOOH \rightleftharpoons O{=}CCOOH + Ce^{3+} + H^+$
16	$2O{=}CCOOH + H_2O \rightleftharpoons O{=}CHCOOH + HCOOH + CO_2$
17	$Br_2 + HCOOH \rightarrow 2Br^- + CO_2 + 2H^+$
18	$2CH(COOH)_2 + H_2O \rightleftharpoons CH_2(COOH)_2 + HOCH(COOH)_2$

La simulazione al calcolatore dell'oscillazione nella reazione di Belousov-Zhabotinsky comporta 18 stadi elementari e 21 specie chimiche diverse. La reazione prende il nome da due chimici sovietici: B. P. Belousov, al quale si deve la sua scoperta, e A. M. Zhabotinsky, che la perfezionò. I materiali di partenza sono tre sostanze inorganiche, gli ioni bromato (BrO_3^-), bromuro (Br^-) e cerosi (Ce^{3+}) oltre a una sostanza organica, l'acido malonico ($CH_2(COOH)_2$). Un mezzo a base di acido solforico fornisce gli ioni idrogeno (H^+). I prodotti della reazione sono anidride carbonica (CO_2), acido formico (HCOOH) e acido bromomalonico ($BrCH(COOH)_2$). Dato che il cerio oscilla fra i due stati di ossidazione cerico (Ce^{4+}) e ceroso (Ce^{3+}), la soluzione vira alternativamente da gialla a incolore. Le reazioni rappresentate in colore comportano solo specie inorganiche e sono meglio conosciute delle reazioni in nero, le quali coinvolgono specie derivate dall'acido malonico. Questo meccanismo è stato messo a punto all'Università dell'Oregon da Richard M. Noyes con il collega Richard J. Field e Endre Körös dell'Università Eötvös di Budapest.

Il sistema chimico bistabile può essere rappresentato da un'analogia meccanica. Una palla che rotola in una buca di potenziale con due punti minimi, *A* e *B*, può fermarsi in uno dei due. *A* e *B* rappresentano pertanto gli stati stazionari stabili del sistema; *C* è uno stato stazionario instabile. Una palla può fermarsi in *C*, ma la più piccola perturbazione la farà scivolare in *A* o in *B*. Qui la variabile è la posizione della palla; il vincolo esterno è rappresentato dalla forma della buca di potenziale.

Il ciclo di isteresi è illustrato in un sistema chimico bistabile, che opera in un reattore a flusso e agitazione continui. Quando la concentrazione X_0 della sostanza X che fluisce nel reattore aumenta, il valore dello stato stazionario Y_{SS} della specie Y nel reattore decresce lentamente lungo la curva in rosso finché, al punto critico A, cade da Y_{A1} a Y_{A2}. Se X_0 viene ulteriormente aumentata, Y_{SS} cade ancora, ma lungo la curva in blu. Se X_0 diminuisce, Y_{SS} ripercorre la curva in blu al di là del punto A finché raggiunge un secondo punto critico B. Qui Y_{SS} balza da Y_{B2} a Y_{B1}. Tra A e B può esistere l'uno o l'altro stato stazionario.

Può verificarsi oscillazione se una adatta sostanza Z viene aggiunta a un sistema chimico bistabile che presenta isteresi. Se Z reagisce con la sostanza Y a bassa velocità per dare la sostanza X, la concentrazione effettiva di quest'ultima *(X_0)* viene modificata in modo che il sistema «veda» un valore di X più elevato di quello fornito dal recipiente *(a sinistra)*. L'effetto di una data quantità di Z su X_0 *(frecce in verde)* è maggiore nel tratto rosso del ciclo di isteresi (dove c'è più Y) rispetto al tratto blu. Immaginiamo che la reazione inizi con un flusso di X uguale a C. Ora aggiungiamo abbastanza Z, una quantità Z_0, così che l'effettiva X_0 sia uguale a C_1 quando il sistema è nel tratto rosso e a C_2 quando è nel tratto blu. In assenza di Z, il flusso $X_0 = C$ darebbe un valore di stato stazionario Y_C. Invece, con l'aggiunta di Z_0 il sistema segue il tratto rosso quando il valore effettivo di X aumenta lentamente per effetto della reazione di Y con Z. Il sistema tenta di raggiungere il valore di C_1, ma nel raggiungere A cade dal tratto rosso e subisce una rapida transizione al tratto blu. A questo punto, il sistema «vede» un valore di $X_0 = C_2$ e così devìa verso sinistra lungo il tratto blu nel tentativo di adattare il nuovo effettivo valore di X_0. Prima che raggiunga lo scopo, a $X_0 = B$ salta di nuovo al tratto rosso e il ciclo è completo *(frecce in giallo)*:

ora X_0 ha nuovamente il valore effettivo C_1, così che il sistema si muove ancora verso destra. La ripetizione di questa sequenza dà origine a un'oscillazione periodica nel valore di Y *(al centro)*. Il diagramma a più fasi *(a destra)* mostra come l'innalzamento di Z_0 può spostare il sistema dalla bistabilità all'oscillazione. Se Z è uguale a 0, il sistema mostra due condizioni di stato stazionario: lo stato «rosso» per un basso valore di X_0 e lo stato «blu» per un valore elevato di X_0. Quando X_0 viene a cadere tra A e B, il sistema è bistabile. Con l'aumentare di Z_0, la regione di bistabilità si stringe finché a un valore critico inizia l'oscillazione.

nee. Il «brodo» ottenuto, quando viene alimentato con appropriate sostanze nutritive, mostra la sua oscillazione nella concentrazione degli ioni idrogeno e della molecola di nicotinammideadenindinucleotide, che nella sua forma ridotta (NADH) trasporta elettroni nel normale ciclo energetico della cellula. Alla Johnson Research Foundation, Britton Chance e collaboratori hanno dimostrato che molti sistemi oscillanti biologici di questo tipo comportano la glicolisi, il processo mediante il quale le cellule producono energia metabolizzando zucchero in assenza di ossigeno. Poiché tali reazioni sono catalizzate da molti enzimi, la loro complessità è davvero scoraggiante.

Nel 1979 due di noi (Epstein e Kustin) hanno deciso che era giunto il momento di tentare di ideare una procedura sistematica per mettere a punto un sistema chimico oscillante. Ma da dove si doveva incominciare? Siamo partiti identificando tre condizioni che sapevamo essere necessarie per l'oscillazione chimica, o che contribuivano ad essa. La prima condizione è che i sistemi chimici possono oscillare solo se sono lontani dall'equilibrio. La seconda è la retroazione *(feedback)*: parte del prodotto di uno stadio della sequenza di reazioni deve esercitare un'influenza sulla sua stessa velocità di formazione. La terza è che il sistema chimico deve mostrare bistabilità. Nelle stesse condizioni esterne, definite vincolanti, il sistema deve poter esistere in due differenti stati stazionari stabili.

Il requisito che il sistema sia mantenuto lontano dall'equilibrio potrebbe essere soddisfatto facendo avvenire la reazione in un apparecchio ben noto agli ingegneri chimici: il reattore a flusso e agitazione continui (CSTR, *continuous-flow stirred-tank reactor)*. Alla metà degli anni settanta, Adolphe Pacault e collaboratori del Centre de Recherches Paul Pascal di Bordeaux hanno adattato il CSTR allo studio di reazioni chimiche oscillanti. Quando, nel 1980, uno di noi (De Kepper) ha lasciato Bordeaux per raggiungere il nostro gruppo alla Brandeis University, è parso naturale costruire un reattore di questo tipo. L'arrivo dall'Ungheria di Orbán, che aveva lavorato a lungo su sistemi oscillanti a base di bromato, ha completato il nostro gruppo.

Un tipo comune di retroazione incontrata nei sistemi viventi è l'autocatalisi, in cui la velocità con la quale una sostanza è prodotta aumenta con la concentrazione. Il concetto si estende a popolazioni di organismi: il numero di organismi aggiunti a una popolazione in un dato intervallo è quasi sempre proporzionale al numero di individui già presenti nella popolazione. (Questa osservazione mette in rilievo l'affermazione, fatta da Thomas Malthus nel 1798, secondo la quale le popolazioni tendono a esaurire le scorte di cibo.) In chimica i sistemi autocatalitici sono rari, ma non straordinari. Già nel 1920 Alfred J. Lotka della Johns Hopkins University dimostrò che un semplice schema, costituito da due reazioni autocatalitiche accoppiate, avrebbe dato luogo a oscillazione. Il meccanismo di Lotka, pur essendosi dimostrato che non era applicabile ad alcuna reazione chimica effettiva, è servito da guida per molti chimici ed è stato di considerevole utilità nel descrivere le oscillazioni di popolazioni di predatori e di prede nei sistemi ecologici.

Consideriamo ora il ruolo della bistabilità, ossia l'esistenza di due differenti stati stazionari stabili in un sistema chimico oscillante. Per stato stazionario si intende una condizione del sistema in cui tutte le variabili, come la concentrazione di ciascuna specie chimica, hanno raggiunto valori costanti. Lo stato stazionario è stabile se può adattarsi a piccoli cambiamenti di una variabile, come l'aggiunta di una goccia di acido, rimanendo tale senza essere trasformato in un nuovo stato. Se la più piccola variazione provoca una transizione a un differente stato del sistema, si dice che lo stato originario è instabile. Per esempio, una piccola palla sul fondo di un recipiente si trova in uno stato stabile. Una palla in equilibrio sul bordo del recipiente è in uno stato instabile. In un reattore a flusso e agitazione continui le condizioni vincolanti esterne che controllano i valori dello stato stazionario sono generalmente la temperatura del bagno termostatico, le concentrazioni dei reagenti chimici nel serbatoio di alimentazione del reattore e le loro velocità di flusso.

Se vengono cambiate le condizioni vincolanti esterne in un sistema chimico bistabile, può verificarsi un fenomeno più particolare, noto come isteresi. Questo fenomeno è comune nel magnetismo, dove assume la forma di un «ciclo di isteresi». Quando un pezzo di ferro è sottoposto a un aumento di forza magnetica, esso raggiunge alla fine la completa saturazione magnetica. Se la forza magnetica si riduce poi a zero, il ferro conserva parte della sua magnetizzazione. Al fine di ridurre anche questa a zero, il campione di ferro va sottoposto a una forza magnetica in senso opposto. Il ciclo completo di magnetizzazione e smagnetizzazione prende l'aspetto di una *S*, spessa al centro e appuntita alle estremità.

Un fenomeno analogo può verificarsi in un sistema chimico bistabile. Nel 1979 un calcolo effettuato da Jacques Boissonade e da uno di noi (De Kepper) al Centre de Recherches Paul Pascal suggerì che le oscillazioni potrebbero essere indotte aggiungendo un altro composto chimico a un sistema capace di isteresi. Il comportamento di un sistema di questo tipo può essere rappresentato in un diagramma a più fasi, il quale mostra che, se la sostanza aggiunta esercita effetti diversi sui due rami degli stati stabili del sistema, è probabile che si sviluppino oscillazioni.

Il nostro piano per ideare un sistema chimico oscillante comprendeva i quattro stadi seguenti. Trovare un sistema autocatalitico, far avvenire la reazione in un reattore a flusso continuo, variare le condizioni fino a trovare una regione di bistabilità e, quindi, introdurre un'altra sostanza capace di influire in modo diverso sui due rami di bistabilità e pertanto di indurre oscillazioni. La nostra ricerca sulla letteratura chimica per le reazioni autocatalitiche ci condusse alla scoperta di molte reazioni interessanti. Due ci sembrarono particolarmente attraenti poiché avevano come intermedio comune lo iodio. Una reazione coinvolgeva gli ioni iodato e l'arsenito, la seconda gli ioni ioduro e il clorito. Come si potrà immaginare, una reazione poteva avere un'auspicabile influenza perturbante sull'altra. La reazione con iodato e arsenito si rivelò ben presto bistabile. Quando venne introdotto il clorito, il sistema composito iniziò quasi immediatamente a oscillare. Il primo sforzo sistematico per ideare un nuovo sistema chimico oscillante fu, quindi, coronato da successo.

Variando i componenti che entrano nel reattore a flusso continuo, noi abbiamo da tempo esteso il sistema oscillante iniziale clorito-iodato-arsenito a un gruppo di due dozzine di sistemi oscillanti correlati, che hanno tutti in comune lo ione clorito. Possiamo tracciare una classificazione dei sistemi oscillanti a base di clorito e stiamo cominciando a vedere in che modo possono essere messi in relazione con i sistemi oscillanti delle famiglie del bromato e dello iodato.

Benché i sistemi bromato siano stati fra i primi a essere scoperti e il loro meccanismo sia meglio conosciuto di quello dei sistemi clorito, la varietà dei sistemi clorito, che abbiamo trovato deliberatamente, è di gran lunga più ampia. Due fattori, uno chimico e l'altro storico, spiegano questo apparente paradosso. Dal punto di vista chimico, ogni famiglia di sistemi chimici oscillanti ha un membro minimo (più semplice) dal quale gli altri possono essere derivati per aggiunta di altre sostanze. Il più semplice sistema oscillante a base di clorito, clorito più ioduro, oscilla entro un'ampia gamma di condizioni. Al contrario, il più semplice sistema oscillante a base di bromato, bromato più bromuro più uno ione metallico, funziona solo entro limiti ben precisi. Infatti, anche dopo che l'esistenza del sistema oscillante più semplice a base di bromato era stata prevista da un ricercatore, Kedma Bar-Eli dell'Università di Tel Aviv, la sua

I principali tipi di sistemi chimici oscillanti sono raggruppati, in un tentativo di classificazione, in tre famiglie principali. Il sistema originario di Belousov-Zhabotinsky e tutte le sue numerose varianti si trovano nel ramo della famiglia del bromato che ha ioni metallici. La sistemazione del sistema bromato nel ramo esente da ioni metallici è opera di Körös e di uno degli autori (Orbán). Tutti i sistemi oscillanti a base di clorito sono stati scoperti dagli autori alla Brandeis University. I sistemi al bromato che comportano riducenti inorganici sono stati messi a punto nella stessa sede in condizioni di flusso continuo. La famiglia dello iodato ha il minor numero di membri. Il primo sistema elencato in essa è la reazione oscillante illustrata nel 1921 da William C. Bray dell'Università della California a Berkeley e ignorata per 50 anni. Il secondo sistema nella famiglia dello iodato è un ibrido tra i sistemi Belousov-Zhabotinsky e Bray. Linee tratteggiate collegano i sistemi che hanno alcuni reagenti in comune. Un riducente è una sostanza che cede elettroni a un'altra; ossidante è una sostanza che li acquista.

ricerca non diede frutti finché due di noi (Orbán ed Epstein) riuscirono a trovarlo seguendo il nostro metodo sistematico basato sul diagramma a più fasi.

Dal punto di vista storico, il primo sistema oscillante a base di bromato fu scoperto prima che entrassero in funzione i reattori a flusso continuo. Sistemi del genere oscillavano facilmente in condizioni di flusso zero. I sistemi a base di clorito, come la grande maggioranza dei nuovi sistemi, mostrano un comportamento periodico solo in condizioni di flusso continuo, poiché il flusso è necessario a mantenerli lontani dall'equilibrio. Un sistema che oscilla in condizioni di flusso zero quasi sempre oscillerà anche in condizioni di flusso, ma raramente avviene il contrario. La maggior parte dei sistemi bromato fu originariamente ricercata in condizioni di flusso zero, mentre quelli clorito furono in origine ricercati solo in condizioni di flusso. Di conseguenza, la varietà di sistemi a base di clorito, che sono stati scoperti, ha presto superato la varietà dei sistemi a base di bromato. (Trovammo successivamente, però, che molti sistemi oscillanti a base di clorito operano anch'essi in condizioni di flusso zero.) Si può dire che l'estrema facilità di oscillazione dei primi sistemi a base di bromato ritardò la ricerca di altri sistemi. Soltanto recentemente i reattori a flusso continuo sono stati impiegati in modo intensivo nella ricerca di nuovi sistemi oscillanti a base di bromato e allo studio di una terza famiglia, finora la più piccola: i sistemi a base di iodato.

Uno degli scopi della progettazione di un nuovo sistema chimico oscillante fu quello di poter disporre di un sistema il cui meccanismo potesse essere decifrato e confrontato con quello delle reazioni oscillanti precedentemente esaminate. L'analisi della famiglia dei sistemi oscillanti a base di clorito è ora ben avviata. Abbiamo completato gli studi delle velocità di numerosi processi componenti e pertanto abbiamo numerose tessere del mosaico; per avere una descrizione completa dello schema oscillante basta ora spostare gruppi di tessere finché combacino tutte insieme nel modo giusto.

1 CLORITO + IODURO → CLORURO + IODIO
2 CLORITO + IODIO → CLORURO + IODATO
3 IODATO + IODURO → IODIO
4 IODIO + ARSENITO → IODURO + ARSENIATO

L'andamento delle oscillazioni nel primo sistema oscillante messo a punto dagli autori può essere compreso qualitativamente considerando l'interazione tra quattro reazioni principali. Il sistema è alimentato con un flusso uniforme e bilanciato di tre soluzioni saline: una di clorito, la seconda di arsenito e l'ultima di iodato. All'interno del reattore a flusso continuo le concentrazioni dello iodio e dello iodato e il rapporto arseniato/arsenito oscillano, come appare dalle tre curve. Al punto *A* le concentrazioni di ioduro e di arsenito sono elevate, mentre quelle dello iodio e dell'arseniato sono basse, così che le reazioni predominanti sono la *1* e la *3*. Per effetto di queste reazioni aumenta bruscamente la concentrazione dello iodio, mentre incomincia a diminuire la concentrazione dello ioduro; il sistema raggiunge il punto *B* e si interrompono le reazioni *1* e *3*. Si svolgono allora le reazioni *2* e *4*, con l'esaurimento dello iodio e la conversione dell'arsenito in arseniato. Il sistema raggiunge il punto *C*, dove l'arsenito e lo ioduro si trovano ai loro livelli minimi, o quasi. L'arsenito viene reimmesso nel reattore e l'arseniato viene scaricato. La reazione *4* rigenera ioduro, mentre consuma iodio e pertanto il sistema ritorna in *A*.

Siamo già in grado di dare un quadro schematico del modo in cui funziona il nostro primo sistema oscillante, quello costituito da soluzioni di sali di potassio o di sodio sotto forma di clorito, iodato e arsenito. Il sistema comporta solo quattro reazioni principali, ciascuna delle quali composta di numerosi stadi elementari. Due di essi comportano la reazione del clorito con lo ioduro o con lo iodio per dare nel primo caso cloruro e iodio e nel secondo cloruro e iodato. Nel terzo stadio lo iodato e lo ioduro reagiscono per formare iodio. Nel quarto lo iodio e l'arsenito formano ioduro e arseniato. Variazioni nelle velocità di queste quattro reazioni danno luogo a oscillazioni periodiche nelle concentrazioni di iodio e iodato e nel rapporto tra arseniato e arsenito *(si veda l'illustrazione in alto)*.

L'oscillazione chimica, benché affascinante di per se stessa, è connessa a numerosi altri fenomeni di almeno pari interesse. Uno è la formazione di strutture spaziali in un mezzo inizialmente omogeneo. Tali strutture si sviluppano, per esempio, in uno dei sistemi oscillanti che abbiamo scoperto, dove i reagenti sono clorito, ioduro e acido malonico. Quando è presente l'amido come indicatore, la soluzione ha inizialmente un colore uniforme rosso violaceo, a causa di un complesso costituito da ioduro, iodio e amido. Con il procedere della reazione appaiono punti bianchi, che si ingrossano fino a formare anelli e serie di anelli concentrici che si annientano quando urtano fra di loro. Un osservatore ha paragonato la comparsa dei punti bianchi nel color rosso circostante alla apparizione improvvisa di stelle. La comparsa di tali forme geometriche in una massa che prima era indifferenziata fa venire subito alla mente il genere di processi che permette alle cellule embrionali negli organismi animali di unirsi in tipi individuali destinati a diventare sangue, cervello o osso.

Un altro fenomeno affine, che ha fatto sorgere una grande controversia fra i

Le oscillazioni periodiche nella reazione clorito-tiosolfato mostrano una varietà di modelli quando viene alterata la velocità di immissione in un reattore a flusso continuo. Ogni ciclo ha una oscillazione ampia e n oscillazioni piccole, dove n varia da 0 a 16. Ciò si verifica per tutti i cicli.

Un'oscillazione aperiodica, o caotica, si osserva talvolta in una stretta gamma di velocità di flusso, tra velocità che danno un comportamento periodico complesso. Una forma di caos consiste in un miscuglio casuale di due tipi di oscillazioni periodiche, una con n e l'altra con $n+1$ piccole oscillazioni per ogni ampia oscillazione. Una possibile spiegazione è che le variazioni incontrollate nella velocità di flusso facciano oscillare a caso il sistema da una velocità che dà oscillazione di tipo n a un'altra che dà oscillazione di tipo $n+1$. L'alternativa, favorita dagli autori, è che l'oscillazione caotica sia intrinseca nella dinamica della reazione chimica a determinate velocità di flusso.

chimici, maggiore di quella sorta un quarto di secolo fa quando fu messa in dubbio la realtà delle reazioni oscillanti, è il «caos chimico». Il termine si riferisce al comportamento di una reazione in cui le concentrazioni non raggiungono valori costanti e neppure oscillano periodicamente, ma piuttosto aumentano e diminuiscono in un modo apparentemente casuale e imprevedibile. Noi abbiamo notato tali fluttuazioni nel sistema BZ e in uno dei nostri nuovi sistemi oscillanti a base di clorito, quando vengono fatti reagire in un reattore a flusso continuo sotto accurato controllo. Essi sono di particolare interesse per i matematici poiché se fossero per natura inerenti alla dinamica della reazione, essi rappresenterebbero esempi fisici di quegli oggetti matematici noti come «attrattori strani». Brevemente e in qualche modo grossolanamente, gli attrattori strani possono comparire quando alcune equazioni sono iterate, o risolte ripetutamente, con il risultato che una iterazione serve da entrata per la successiva. In un sistema di equazioni con un attrattore strano il percorso segnato da soluzioni successive sembra variare in modo imprevedibile da un ciclo all'altro.

Chimici scettici hanno ipotizzato che il caos chimico non rappresenti niente di più che fluttuazioni sperimentali incontrollate, per esempio nella temperatura o nella velocità di flusso, che spingono a caso un sistema da un modo di oscillazione periodica complessa a un altro. Sebbene un giudizio definitivo non sia stato ancora dato, un recente e accurato lavoro sperimentale e teorico di Harry L. Swinney e Jack Turner dell'Università del Texas a Austin, di Jean Claude Roux, Christian Vidal e collaboratori del Centre de Recherches Paul Pascal e di J. L. Hudson dell'Università della Virginia propone che il caos chimico sia un fenomeno genuino inerente alla dinamica di alcuni sistemi chimici oscillanti. Le implicazioni di questo caos intrinseco non sono ancora conosciute. Potrebbe darsi, come Henry Adams osservò una volta, che «il caos spesso genera la vita, quando l'ordine genera abitudine»?

Ora che i sistemi chimici oscillanti hanno dimostrato non solo di essere compatibili con le leggi della natura, ma anche di poter essere ideati e descritti sistematicamente, incominciano ad avere un ruolo sempre più importante in molti settori scientifici. Il passaggio da uno stato a un altro, che caratterizza una reazione oscillante, può serbare la chiave per capire i processi di regolazione nelle cellule viventi, come i meccanismi che innescano o arrestano la copiatura di un filamento di DNA o la contrazione di un muscolo. Le stesse forze che creano anelli e strati colorati in sistemi chimici oscillanti non sottoposti a rimescolamento possono anche essere responsabili della separazione degli anelli di Saturno e delle striature periodiche in alcune formazioni rocciose, spiegate con difficoltà dai processi geologici convenzionali.

Sembra inoltre plausibile che molte e

In un sistema chimico oscillante a base di clorito, ioduro e acido malonico, scoperto dagli autori, si sviluppa una struttura spaziale. La soluzione, posta in una piastra di vetro poco profonda, è inizialmente di un colore rosso brunastro uniforme per la formazione di un complesso tra iodio, ioduro e amido, usato come indicatore. Al procedere della reazione appaiono punti bianchi, che si ingrossano fino a formare anelli e serie di anelli concentrici, che si annullano quando si urtano. L'ultima fotografia è stata scattata circa 90 secondi dopo l'inizio della reazione.

forse tutte le reazioni catalitiche che sono fondamentali per l'industria chimica procedano in modo oscillante. Il catalizzatore può alternarsi tra due forme in un intervallo di tempo troppo breve perché sia rilevato dai normali processi di analisi. Lo studio di reazioni oscillanti può fornire un terreno di prova per teorie sulla catalisi e, più in generale, sulla dinamica delle reazioni chimiche. Qualunque modello possa prevedere con successo il complesso comportamento di sistemi oscillanti dovrebbe essere in grado di trattare senza difficoltà i fenomeni più semplici osservati in reazioni meno strane. Aumentando il numero dei sistemi chimici oscillanti trovati o ideati deliberatamente, si può sperare di trovare la chiave per penetrare in una varietà di fenomeni, compresi quegli «orologi» e «caroselli» che operano all'interno del nostro corpo.

Caos quantistico?

Il tempo che farà domani, come il decadimento radioattivo di un atomo, può essere descritto soltanto in termini probabilistici. Due fattori molto diversi sono all'origine dell'indeterminazione che non consente di prevedere con esattezza il corso di questi eventi. La natura caotica, ossia casuale, di un sistema macroscopico come l'atmosfera deriva in gran parte dagli innumerevoli e sconvolgenti effetti che possono essere prodotti nel corso del tempo da piccole fluttuazioni nelle variabili di equazioni meccaniche deterministiche. Al contrario, la descrizione dei fenomeni a livello subatomico si basa sulle equazioni probabilistiche della meccanica quantistica; la casualità ha origine da un limite intrinseco alla precisione con cui è possibile misurare certe variabili.

È possibile che queste due fonti di casualità abbiano un comune denominatore? Se così non fosse, nella struttura della fisica moderna vi sarebbe una scoraggiante separazione, a livello fondamentale, fra fenomeni microscopici e fenomeni macroscopici. In effetti, recenti analisi al calcolatore eseguite da un gruppo di fisici italiani e sovietici implicano che il caos tanto prevalente nel mondo macroscopico sia estraneo all'ambito della meccanica quantistica. Queste ricerche inducono a chiedersi fino a che punto la meccanica quantistica sia coerente con la teoria del caos e, secondo alcuni scienziati, mettono perfino in dubbio la compatibilità della meccanica quantistica con la fisica dell'universo in generale.

Un mezzo adeguato per verificare la concordanza (o la non concordanza) delle due teorie è costituito dai sistemi «semiclassici», ossia quelli in cui le interazioni fra particelle e campi possono essere trattate sia in maniera classica sia in maniera quantistica. Un sistema di questo tipo - la ionizzazione di un atomo di idrogeno altamente eccitato da parte di microonde - è stato analizzato da Giulio Casati dell'Università degli Studi di Milano, Boris V. Čirikov e Dima L. Šepeljanskij dell'Istituto di fisica nucleare di Novosibirsk e Italo Guarneri dell'Università di Pavia. Generalmente si ammette che un atomo di idrogeno si ionizza quando il suo unico elettrone assorbe un singolo fotone di energia superiore a una certa soglia. Ad attrarre l'attenzione dei ricercatori è stata la prova sperimentale che la ionizzazione può avvenire anche a frequenze, o energie, molto inferiori al valore di soglia relativo all'assorbimento di un singolo fotone.

Come riferiscono i ricercatori citati in «Physical Review Letters» 57, n. 7, agosto 1986, il problema può essere affrontato da un punto di vista classico analizzando la ionizzazione come un processo caotico o, in alternativa, può essere formulato in termini quantistici. Gli studiosi hanno seguito entrambe le impostazioni teoriche e hanno scoperto che in certe condizioni sia l'una sia l'altra danno una probabilità di ionizzazione elevata a frequenze inferiori alla soglia di assorbimento di un singolo fotone. Nella formulazione quantistica, tuttavia, la probabilità di ionizzazione si annulla bruscamente a una frequenza (ancora inferiore al valore di soglia) alla quale la probabilità classica è maggiore di zero.

Questi risultati, come osserva in «Nature» vol. 325, 1 gennaio 1987 Joseph Ford del Georgia Institute of Technology, sono interessanti per tre aspetti. In primo luogo, essi consentono di verificare teoricamente che si può avere ionizzazione al di sotto della soglia relativa all'assorbimento di un singolo fotone. Di fatto i calcoli di Casati e colleghi mostrano che in condizioni appropriate questo tipo di ionizzazione è molto più probabile del processo a fotone singolo. Ford ipotizza che questo effetto, qualora fosse reale, potrebbe essere sfruttato per migliorare l'efficienza di dispositivi elettronici che attualmente si basano sul processo a fotone singolo.

In secondo luogo, l'intervallo di frequenze nel quale l'elettrone può ionizzarsi nella formulazione classica, ma non in quella quantistica, offre un'occasione per verificare la validità delle due impostazioni teoriche. Se precise misurazioni di laboratorio indicassero che si può avere ionizzazione in questo intervallo di frequenze si renderebbe forse necessaria una revisione delle equazioni della meccanica quantistica. Infine il sistema aggiunge un altro paradosso al mondo già paradossale delle particelle subatomiche: un elettrone può essere descritto come caotico o non caotico a seconda che si consideri il sistema da un punto di vista classico o da uno quantistico.

Altri fisici concordano con Ford che le prime due questioni dovrebbero essere risolte sottoponendo la previsione teorica alla verifica sperimentale, mentre il paradosso potrebbe essere eliminato con un semplice chiarimento semantico. Norman H. Packard dell'Università dell'Illinois a Urbana-Champaign fa notare che la definizione classica di caos è attinente all'evoluzione di un sistema nel tempo; tuttavia «è impossibile seguire l'evoluzione temporale tramite misurazioni quantistiche». Sebbene le equazioni dipendenti dal tempo che descrivono un sistema quantistico siano ben definite, lo stato fisico corrispondente del sistema non lo è; le equazioni definiscono essenzialmente il sistema come una sovrapposizione di tutti i possibili stati fisici. Ciò che si osserva quando si compie una misurazione di un sistema quantistico è solo uno dei molti possibili stati, «congelatosi» fortuitamente nella realtà.

Secondo Ford, tuttavia, il paradosso non è conciliabile. Egli prevede che, se i risultati sperimentali gli daranno ragione, la visione quantistica dovrà essere respinta in favore di una teoria universale del caos.

Il caos e i limiti teorici del calcolo

Settembre 1988
Numero 241
Anno XXI
Volume XLI

La casualità in aritmetica

Non c'è un metodo per stabilire se ogni membro di una famiglia di equazioni algebriche possiede un numero di soluzioni finito o infinito: le risposte variano con casualità e sfuggono al ragionamento matematico

di Gregory J. Chaitin

Che cosa potrebbe essere più certo del fatto che 2 più 2 è uguale a 4? Fin dal tempo dell'antica Grecia, i matematici hanno avuto la convinzione che poche cose - se pure ce ne sono - siano indubitabili come un teorema dimostrato. In effetti, gli enunciati matematici di cui si può dimostrare la verità sono stati spesso ritenuti un fondamento più solido, per un sistema di pensiero, di qualsiasi massima morale o anche della stessa oggettività fisica. Il matematico e filosofo tedesco del XVII secolo Gottfried Wilhelm Leibniz giunse addirittura a immaginare un ragionamento in forma di calcolo grazie al quale si potessero un giorno risolvere tutte le dispute seguendo l'invito: «Signori, calcoliamo!» All'inizio di questo secolo, la logica simbolica aveva compiuto tali progressi da indurre il matematico tedesco David Hilbert ad affermare che in linea di principio tutte le asserzioni matematiche sono decidibili; fiducioso si accinse così a codificare, una volta per tutte, le regole del ragionamento matematico.

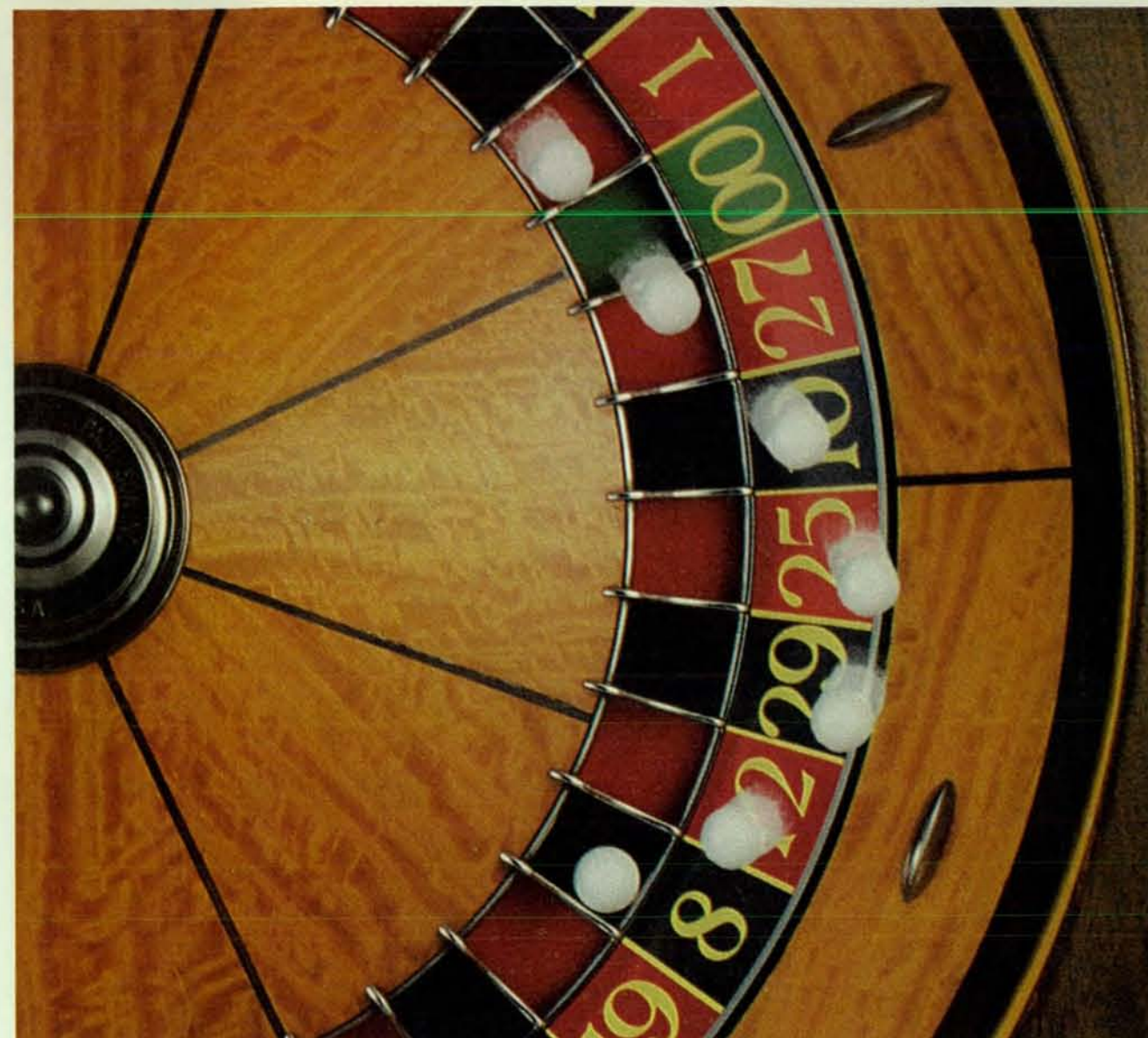
Tale lodevole ottimismo si infranse contro le profonde e sorprendenti scoperte di Kurt Gödel e Alan M. Turing negli anni trenta. Gödel dimostrò che nessun insieme finito di assiomi e regole di deduzione poteva racchiudere tutte le proprietà matematiche degli interi positivi. Turing, in seguito, espresse in una forma più accessibile l'ingegnosa e complicata dimostrazione di Gödel. Egli dimostrò l'equivalenza tra il teorema di incompletezza di Gödel e l'asserzione che non esiste un metodo generale per poter decidere sempre se un programma per calcolatore si fermerà, cioè se esso smetterà mai di girare sul calcolatore. Ovviamente, se un certo programma porta il calcolatore ad arrestarsi, è facile dimostrarlo: basta farlo girare. Il difficile sta nel dimostrare, per un programma arbitrario, che non si fermerà mai.

Di recente ho compiuto un passo avanti lungo la via segnata da Gödel e Turing. Traducendo uno specifico programma per calcolatore in un'equazione algebrica di un tipo noto fin dal tempo dell'antica Grecia, ho dimostrato che esiste una casualità in quella branca della matematica pura nota come teoria dei numeri. Il mio lavoro dimostra che - per usare una metafora di Einstein - Dio talvolta gioca a dadi con i numeri interi!

Questo risultato, che si inserisce nel campo della cosiddetta teoria algoritmica dell'informazione, non deve indurre al pessimismo; non porterà la matematica all'anarchia. (Tant'è vero che la maggior parte dei matematici continua a lavorare come prima.) Il suo significato risiede nella considerazione che in certe situazioni si dovrebbero forse applicare leggi matematiche di tipo diverso: leggi statistiche. Così come è impossibile prevedere il momento esatto in cui un singolo atomo subisce un decadimento radioattivo, allo stesso modo la matematica è talora incapace di rispondere a certe domande. Ciò non toglie che i fisici possano fare affidabili previsioni di media su grandi insiemi di atomi. I matematici, forse, dovrebbero in certi casi limitarsi a seguire un metodo di questo genere.

Il mio lavoro è la naturale continuazione di quello di Turing, ma mentre Turing si poneva il problema di stabilire se un programma qualsiasi si sarebbe mai fermato, io mi pongo quello della probabilità che un programma, scelto in modo del tutto casuale, smetta di girare su un qualsiasi calcolatore di uso generale. Che cosa intendo per «scelto in modo del tutto casuale»? Poiché qualsiasi programma è sostanzialmente riducibile a una successione di bit (ciascuno dei quali può assumere il valore 0 o il valore 1) che vengono «letti» e «interpretati» dall'hardware del calcolatore, ritengo che un programma completamente casuale di n bit potrebbe benissimo essere ottenuto con n lanci di moneta in cui a ogni «testa» corrisponda uno 0 e a ogni «croce» corrisponda un 1 o viceversa.

La probabilità che tale programma casuale si fermi, probabilità che ho chiamato omega (Ω), si può esprimere come un numero reale compreso tra 0 e 1. (Per $\Omega=0$ nessun programma casuale si fermerebbe mai e per $\Omega=1$ tutti i programmi casuali si fermerebbero. Per un calcolatore di uso generale nessuno di questi casi è in realtà possibile.) Essendo un numero reale, Ω si può esprimere completamente solo come successione infinita di cifre. Questa, espressa in base 2, risulterebbe una sequenza infinita di 0 e 1.

Forse la caratteristica più interessante di Ω è quella di essere algoritmicamente casuale: non può essere «compressa» in un programma (considerato come successione di bit) più breve di se stessa. Questa definizione di casualità ha un

L'imprevedibilità è un concetto abituale nei casinò ed è il fattore che consente ai loro gestori di ottenere un profitto in giochi come la roulette. L'autore sostiene che forse anche i matematici dovrebbero adattarsi all'idea che una certa imprevedibilità sia inerente a ciò di cui si occupano. Essa sorge dalla casualità che esiste nelle soluzioni di problemi piuttosto semplici della teoria dei numeri.

ruolo centrale nella teoria algoritmica dell'informazione e venne formulata in modo indipendente da A. N. Kolmogorov e da me verso la metà degli anni sessanta (in seguito ho dovuto correggerla).

L'idea che sta alla base della definizione è molto semplice. Alcune successioni di cifre possono essere compresse in programmi molto più brevi di quanto esse siano, perché seguono uno schema o una regola. Così, una successione di 200 bit che abbia la forma 0101010101... può essere compressa descrivendola come «100 ripetizioni di 01». Successioni di questo genere non sono certo casuali. Una successione di 200 bit generata lanciando una moneta, invece, non può essere compressa, perché in generale non esiste una regolarità nella successione di 0 e 1: la sequenza è completamente casuale.

Di tutte le possibili sequenze di bit, molte non sono comprimibili e sono quindi casuali. Dato che una successione di bit può essere considerata come una rappresentazione in base 2 di qualche numero reale (se si ammettono successioni infinite), ne consegue che molti numeri reali sono casuali. Non è difficile dimostrare che un numero algoritmicamente casuale, come Ω, mostra le proprietà statistiche normalmente associate alla casualità. Tra queste vi è la normalità: ogni cifra possibile appare con uguale frequenza nel numero. In una rappresentazione in base 2 ciò significa che, con l'approssimarsi all'infinito del numero di cifre di Ω, 0 e 1 appaiono rispettivamente il 50 per cento di volte.

Una condizione tecnica fondamentale che deve essere stipulata perché Ω abbia senso è che un programma in ingresso deve essere autodelimitante: la sua lunghezza totale (in bit) deve essere data all'interno del programma stesso. (Questo problema apparentemente minore, e che invece ha paralizzato per circa un decennio ogni progresso nel settore, è all'origine della ridefinizione di casualità algoritmica.) I linguaggi di programmazione reali sono autodelimitanti, in quanto dispongono di costrutti per iniziare e terminare un programma. Questi costrutti consentono a un programma di contenere sottoprogrammi ben definiti, i quali possono avere annidati nel loro interno altri sottoprogrammi. Dato che un programma autodelimitante è co-

struito tramite la concatenazione e l'annidamento di sottoprogrammi autodelimitanti, un programma è sintatticamente completo solo quando viene chiuso l'ultimo sottoprogramma aperto. In sostanza, i costrutti di inizio e di fine per i programmi e i sottoprogrammi funzionano come le parentesi sinistra e destra delle espressioni matematiche.

Se i programmi non fossero autodelimitanti, non potrebbero essere costruiti a partire da sottoprogrammi, e sommando le probabilità di fermata per tutti i programmi si arriverebbe a un numero infinito. Se si considerano solo programmi autodelimitanti, non solo Ω è limitato all'intervallo tra 0 e 1, ma può anche essere calcolato esplicitamente per approssimazioni successive. Vale a dire, è possibile calcolare una successione infinita di numeri razionali (che si possono esprimere in termini di successione finita di bit), ciascuno dei quali è più vicino del precedente al valore effettivo di Ω.

Un modo per effettuare questa operazione consiste nel calcolare sistematicamente Ω_n per valori crescenti di n, dove Ω_n è la probabilità che un programma completamente casuale di dimensioni fino a n bit si fermi entro n secondi se viene fatto girare su un dato calcolatore. Essendoci 2^k possibili programmi lunghi k bit, in linea di principio si può calcolare Ω_n determinando, per ogni valore di k compreso tra 1 e n, quanti dei possibili programmi si fermino effettivamente entro n secondi, moltiplicando quel numero per 2^{-k} e poi sommando tutti i prodotti. Ossia, ciascun programma di k bit che si fermi contribuisce per 2^{-k} a Ω_n; i programmi che non si fermano danno un contributo pari a 0.

Se si potesse miracolosamente avere il valore di Ω con k bit di precisione, si potrebbe calcolare una successione di Ω_n fino a raggiungere un valore che uguagli quello di Ω. A questo punto si conoscerebbero tutti i programmi di dimensioni inferiori a k bit che si fermano; in sostanza, si sarebbe risolto il problema di Turing della fermata per tutti i programmi di dimensione inferiore a k bit. Naturalmente, il tempo richiesto dal calcolo, per valori di k ragionevoli, sarebbe enorme.

Finora ho fatto riferimento solo ai calcolatori e ai loro programmi per parlare del problema della fermata, ma questo problema ha assunto una nuova dimensione alla luce del lavoro di J. P. Jones, dell'Università di Calgary, e di Y. V. Matijasevič, dell'Istituto V. A. Steklov di matematica di Leningrado. Il loro lavoro offre un metodo per affrontare il problema sotto forma di asserzioni su particolari equazioni diofantee. Queste equazioni algebriche, che coinvolgono solo moltiplicazione, somma ed elevamento a potenza di numeri interi, prendono il nome da Diofanto di Alessandria, matematico greco del III secolo.

Più specificamente, applicando il metodo di Jones e Matijasevič si può equiparare l'enunciato secondo cui un certo programma non si ferma all'asserzione che un'equazione appartenente a una particolare famiglia di equazioni diofantee non ha alcuna soluzione nei numeri interi. Come nel caso della versione originale del problema della fermata per i calcolatori, è facile dimostrare, quando c'è, che una soluzione esiste: basta inserire i numeri corretti e verificare che i numeri risultanti alla destra e alla sinistra del segno di uguaglianza siano effettivamente uguali. Molto più difficile è dimostrare che non vi sono assolutamente soluzioni quando questo è il caso.

La famiglia di equazioni è costruita a partire da un'equazione di base che contiene una particolare variabile k, chiamata il parametro, che assume i valori 1, 2, 3 e così via (*si veda l'illustrazione a pagina 46, in alto*). Esiste quindi una famiglia infinitamente grande di equazioni (una per ciascun valore di k) che si può generare partendo da un'equazione di base per ciascun membro di una «famiglia» di programmi. L'asserzione matematica che l'equazione diofantea con parametro k non ha alcuna soluzione corrisponde all'asserzione che il k-esimo programma per calcolatore non si ferma mai. Se invece il k-esimo programma si ferma, l'equazione ha esattamente una soluzione. In un certo senso, la verità o la falsità di asserzioni di questo tipo è matematicamente incerta, in quanto varia in modo imprevedibile a mano a mano che k assume valori differenti.

Il mio modo di affrontare la questione dell'imprevedibilità in matematica è analogo, ma contempla un grado molto maggiore di casualità. Invece di «aritmetizzare» sotto forma di famiglia di equazioni diofantee i programmi per calcolatori che forse si fermano e forse no, io applico il metodo di Jones e Matijasevič per aritmetizzare un singolo programma al fine di calcolare il k-esimo bit di Ω_n.

Il metodo si basa su una curiosa proprietà della parità dei coefficienti binomiali (se sono pari o dispari) notata da Édouard A. Lucas un secolo fa, e finora mai considerata nel suo giusto valore. I coefficienti binomiali sono i moltiplicandi delle potenze di x generati dalla costruzione di quello che è noto come triangolo di Pascal.

Il teorema di Lucas afferma che il coefficiente di x^k nell'espansione di $(x+1)^n$ è dispari solo se ciascuna cifra della rappresentazione in base 2 del numero k è inferiore o uguale alla corrispondente cifra della rappresentazione in base 2 di n (partendo da destra e leggendo verso sinistra). Per dirla in termini un po' più semplici, il coefficiente per x^k in un'espansione di $(x+1)^n$ è dispari se per ogni bit di k che sia un 1, anche il corrispondente bit di n è un 1; in caso contrario è pari. Per esempio, il coefficiente di x^2 nell'espansione binomiale di $(x+1)^4$ è 6, che è pari. Quindi l'1 nella rappresentazione in base 2 di 2(10) non corrisponde a un 1 che occupi la stessa posizione nella rappresentazione in base 2 di 4(100).

Pur essendo concettualmente semplice ed elegante, l'aritmetizzazione richiede un sostanzioso impegno di programmazione. Mi sembrava, in ogni caso, che la cosa potesse risultare divertente e ho elaborato quindi un programma «compilatore» per produrre equazioni da programmi per una macchina a registri. Una macchina a registri è un calcolatore formato da un piccolo insieme di registri per immagazzinare numeri di qualsiasi grandezza. Si tratta, naturalmente, di un'astrazione, dato che i registri di qualsiasi calcolatore reale hanno una capacità limitata.

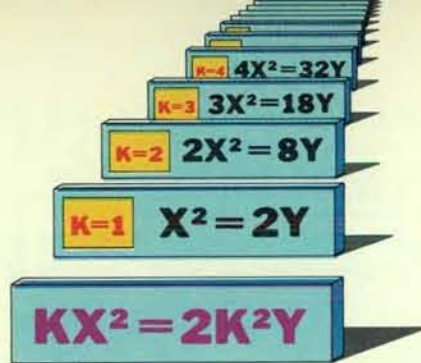
Fornendo in ingresso a un calcolatore reale programmato con il compilatore un programma per macchina a registri che esegua istruzioni scritte nel linguaggio LISP, si ottiene in uscita, entro pochi minuti, un'equazione lunga circa 200 pagine e con circa 17 000 variabili intere non negative. Posso così derivare un'equazione diofantea con un parametro k che codifichi il k-esimo bit di Ω_n semplicemente inserendo nell'equazione di 200 pagine un programma LISP (in forma binaria) per calcolare il k-esimo bit di Ω_n. Per qualsiasi coppia data di valori di k e di n, l'equazione diofantea ha esattamente una soluzione se il k-esimo bit di Ω_n è un 1, e non ha alcuna soluzione se il k-esimo bit di Ω_n è uno 0.

Dato che questo vale per qualsiasi

DAVID HILBERT (1900):
"Ogni problema matematico definito deve essere necessariamente suscettibile di una chiarificazione esatta, o nella forma di una risposta effettiva alla domanda sollevata, o attraverso la dimostrazione dell'impossibilità della sua soluzione e quindi del necessario fallimento di tutti i tentativi... Per quanto inaccessibili possano sembrarci questi problemi e per quanto si possa rimanere scoraggiati di fronte a essi, abbiamo, nondimeno, la ferma convinzione che la loro soluzione deve conseguire da un numero finito di processi puramente logici...
"Sentiamo in noi un perpetuo richiamo: C'è il problema. Cerca la soluzione. Puoi trovarla con la pura ragione..."

David Hilbert, qui ritratto all'età di circa 50 anni, sostenne la decidibilità in linea di principio di tutti i problemi matematici; ossia, che sarebbe stato possibile dimostrare o confutare tutti gli enunciati matematici sulla base di un ristretto numero di assiomi e regole.

KURT GÖDEL (1944):
"Si è visto che la soluzione di certi problemi aritmetici richiede l'uso di assunti che sostanzialmente trascendono l'aritmetica... Naturalmente, in queste circostanze la matematica può perdere una buona dose della sua 'certezza assoluta', ma, sotto la spinta della moderna critica dei fondamenti, questo è già in larga misura avvenuto..."

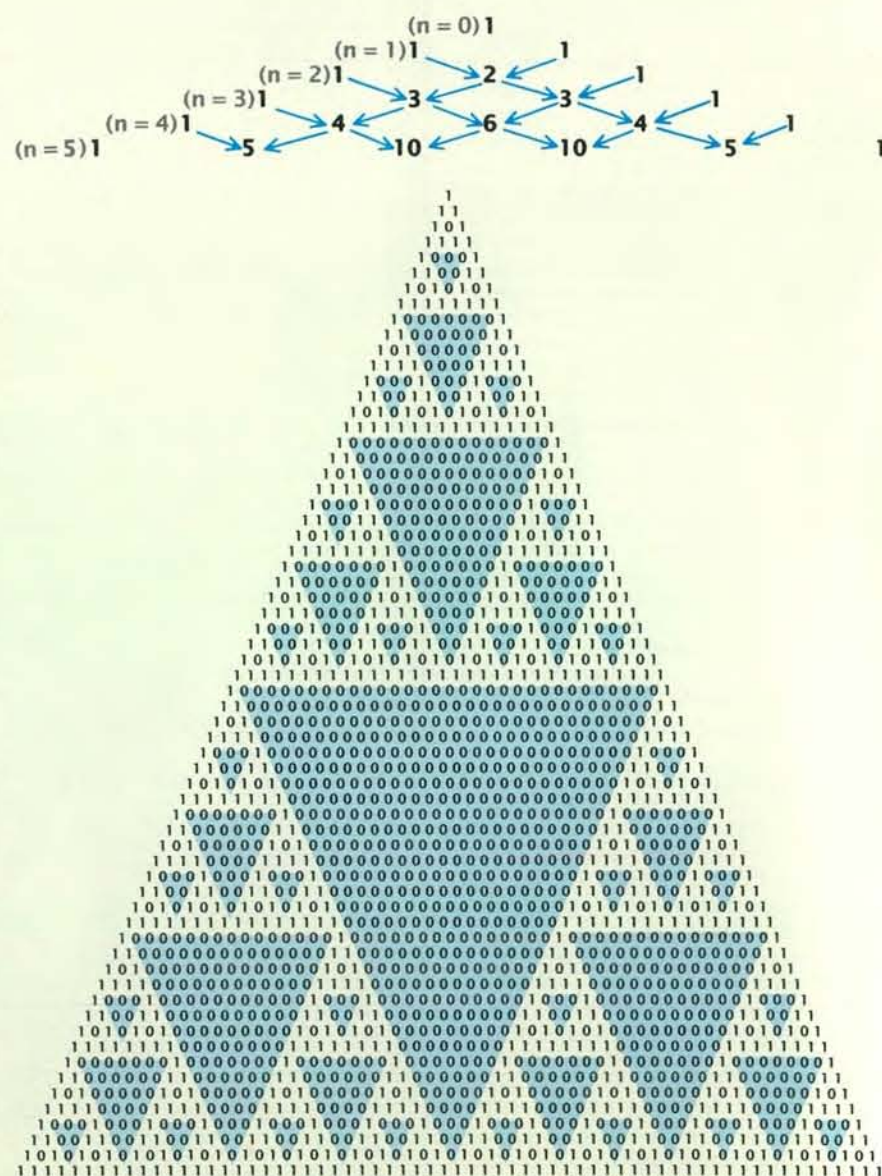
L'indecidibilità essenziale dell'insieme delle questioni matematiche è stata dimostrata da Kurt Gödel, qui ritratto all'età di circa 50 anni in una fotografia di Arnold Newman. Gödel pubblicò la sua dimostrazione nel 1931, quando aveva 25 anni e Hilbert ne aveva 70.

Risulta possibile dare origine a «famiglie» infinitamente grandi di equazioni assegnando come valori numeri interi al parametro k di un'equazione di base nota.

coppia di valori di k e n, in linea di principio si può mantenere fisso k e aumentare sistematicamente il valore di n senza limite, calcolando il k-esimo bit di Ω_n per ciascun valore di n. Per valori piccoli di n, il k-esimo bit di Ω_n fluttuerà in modo irregolare tra 0 e 1. Alla fine, però, si stabilizzerà su 0 o su 1, in quanto per valori molto grandi di n sarà uguale al k-esimo bit di Ω, che è immutabile. Quindi l'equazione diofantea ha effettivamente un numero infinito di soluzioni per un particolare valore del suo parametro k, se il k-esimo bit di Ω risulta essere un 1; mentre, per ragioni analoghe, ha solo un numero finito di soluzioni se il k-esimo bit di Ω risulta essere uno 0. Così, invece di considerare se un'equazione diofantea ha soluzioni per ciascun valore del parametro k, io chiedo se ha un numero infinito di soluzioni.

Potrebbe sembrare che ci sia ben poco da ricavare a chiedersi se ci sono infinite soluzioni invece di chiedersi se ce ne siano alcune; in realtà c'è una differenza fondamentale: le risposte alla mia domanda sono indipendenti dal punto di vista logico. Due asserzioni matematiche sono logicamente indipendenti se è impossibile derivare una dall'altra, cioè se nessuna delle due è una conseguenza logica dell'altra. Questa nozione di indipendenza può di solito essere distinta da quella applicata in statistica, in cui due eventi casuali sono detti indipendenti se l'esito dell'uno non influisce sull'esito dell'altro. Per esempio, il risultato del lancio di una moneta non influisce in alcun modo sul risultato del lancio successivo: i risultati sono statisticamente indipendenti.

Nel mio approccio vengono utilizzate entrambe le nozioni di indipendenza. La risposta al mio quesito per un certo valore di k è logicamente indipendente dalla risposta per un altro valore di k e la ragione di ciò è che i singoli bit di Ω, che determinano la risposta, sono statisticamente indipendenti.

Per quanto sia facile dimostrare che per circa metà dei valori di k il numero di soluzioni è finito e per l'altra metà è infinito, non si può in alcun modo comprimere le risposte in una formula o in un insieme di regole; sono come i risultati dei lanci di una moneta. Dato che Ω è algoritmicamente casuale, anche conoscere le risposte per 1000 valori di k non aiuterebbe a dare la risposta corretta per un altro valore di k. Un matematico, per decidere se una certa equazione ha un numero finito o infinito di soluzioni, non potrebbe fare meglio di uno scommettitore. Quali che siano gli assiomi e le dimostrazioni eventualmente applicabili per trovare la risposta valida per l'equazione diofantea con un certo valore di k, essi risulterebbero inapplicabili per la stessa equazione con un altro valore di k.

In questo caso, il ragionamento matematico non è di alcun aiuto, dato che non esistono connessioni logiche tra le equazioni diofantee così generate. Per quanto si sia perspicace o siano lunghe le dimostrazioni e complessi gli assiomi, la serie infinita di proposizioni che affermano se il numero di soluzioni delle equazioni diofantee è finito o infinito risulterà rapidamente inefficace con il crescere di k. Casualità e imprevedibilità si hanno già nelle branche elementari della teoria dei numeri che abbiano a che fare con le equazioni diofantee.

Il triangolo di Pascal (*in alto*) è un metodo per calcolare i coefficienti nell'espansione di espressioni della forma $(x+1)^n$. Si inizia con un triangolo di soli 1 e si calcolano i valori di ciascuna riga successiva sommando numeri adiacenti e aggiungendo degli 1 a ciascuna estremità della riga. In questo modo si può stabilire, per esempio, che $(x+1)^4 = 1x^4+4x^3+6x^2+4x^1+1x^0$. Il triangolo può essere trasformato in un'attraente configurazione frattale se si sostituiscono i coefficienti dispari con un 1 e i coefficienti pari con uno 0. La configurazione illustra in forma grafica una proprietà dei coefficienti che trova applicazione nell'«aritmetizzazione» in equazioni algebriche di programmi per calcolatore.

In che modo il teorema di Gödel, il problema della fermata di Turing e il mio lavoro hanno influenzato la matematica? La realtà è che la maggior parte dei matematici non hanno dato loro peso. Naturalmente, in linea di principio concordano sul fatto che qualsiasi insieme finito di assiomi è incompleto, ma in pratica la cosa li lascia indifferenti perché non la ritengono pertinente al loro lavoro. Purtroppo, però, questo non è sempre vero. Anche se il teorema originale di Gödel sembrava coinvolgere solo proposizioni matematiche poco consuete che difficilmente avrebbero avuto un interesse pratico, la teoria algoritmica dell'informazione ha dimostrato che l'incompletezza e la casualità sono naturali e pervasive. Questo mi fa pensare che forse andrebbe presa più seriamente in considerazione la possibilità di cercare nuovi assiomi validi per i numeri interi.

Il fatto che molti problemi matematici siano rimasti irrisolti per centinaia e addirittura per migliaia di anni tende a rafforzare la mia tesi. I matematici sono tenacemente convinti che la mancata soluzione di questi problemi dipenda unicamente da loro stessi, ma la difficoltà non potrebbe stare invece nell'incompletezza dei loro assiomi? Per esempio, la questione relativa all'esistenza o meno di numeri dispari perfetti ha eluso una risposta fin dai tempi dell'antica Grecia. (Un numero perfetto è un numero che risulta esattamente dalla somma dei suoi divisori, escluso se stesso. Quindi 6 è un numero perfetto, in quanto è uguale a 1 più 2 più 3.) Non potrebbe darsi che l'enunciato «Non esistono numeri dispari perfetti» non sia dimostrabile? Se così fosse, forse i matematici farebbero meglio ad accettarlo come assioma.

Questa proposta può sembrare ridicola a molti matematici, ma forse a un fisico o a un biologo non apparirebbe assurda. Per chi lavora nelle scienze empiriche, è l'utilità di un'ipotesi, e non necessariamente la sua «verità autoevidente», il criterio chiave con cui giudicare se tale ipotesi possa essere posta a fondamento di una teoria. Se numerose congetture possono essere risolte facendo appello a una certa ipotesi, gli scienziati empirici prendono quell'ipotesi sul serio. (La non esistenza di numeri dispari perfetti non sembra avere significative implicazioni, quindi con questo criterio non sarebbe un assioma utile.)

In realtà, in alcuni casi i matematici hanno già assunto a fondamento del proprio lavoro congetture non dimostrate, ma utili. Per esempio, la cosiddetta ipotesi di Riemann - mai dimostrata - è spesso accettata come vera, perché su essa si basano molti altri importanti teoremi. L'ipotesi, inoltre, è stata sottoposta a verifica empirica con i più potenti calcolatori, e nessuno di essi è arrivato a un solo controesempio. Gli stessi programmi per calcolatore (che, come ho mostrato, sono equivalenti a enunciati matematici) vengono controllati in questo modo, cioè con la verifica di un gran numero di casi sperimentali piuttosto che con una rigorosa dimostrazione matematica.

PROGRAMMA PER MACCHINA A REGISTRI CON ISTRUZIONI IN LISP (300 RIGHE)

PROGRAMMA IN LISP PER CALCOLARE IL k-esimo BIT DI Ω_n (150 RIGHE)

COMPILATORE (700 RIGHE)

EQUAZIONE DIOFANTEA (17 000 VARIABILI)

```
2+(q)**2 + (1+q*i)**2+(i+q**time)**2 + (r1)**2+(L1)**2 + (s1)*
*2+(i)**2 + (t1)**2+(2**s1)**2 + ((1+t1)**s1)**2+(v1*t1**(r1+1
)+u1*t1**r1+w1)**2 + (w1+x1+1)**2+(t1**r1)**2 + (u1+y1+1)**2+(
t1)**2 + (u1)**2+(2*z1+1)**2 + (r2)**2+(L2)**2 + (s2)**2+(i)**
2 + (t2)**2+(2**s2)**2 + ((1+t2)**s2)**2+(v2*t2**(r2+1)+u2*t2*
*r2+w2)**2 + (w2+x2+1)**2+(t2**r2)**2 + (u2+y2+1)**2+(t2)**2 +
 (u2)**2+(2*z2+1)**2 + (r3)**2+(L3)**2 + (s3)**2+(i)**2 + (t3)
**2+(2**s3)**2 + ((1+t3)**s3)**2+(v3*t3**(r3+1)+u3*t3**r3+w3)*
*2 + (w3+x3+1)**2+(t3**r3)**2 + (u3+y3+1)**2+(t3)**2 + (u3)**2
+(2*z3+1)**2 + (r4)**2+(L4)**2 + (s4)**2+(i)**2 + (t4)**2+(2**
s4)**2 + ((1+t4)**s4)**2+(v4*t4**(r4+1)+u4*t4**r4+w4)**2 + (w4
+x4+1)**2+(t4**r4)**2 + (u4+y4+1)**2+(t4)**2 + (u4)**2+(2*z4+1
)**2 + (r5)**2+(L5)**2 + (s5)**2+(i)**2 + (t5)**2+(2**s5)**2 +
 ((1+t5)**s5)**2+(v5*t5**(r5+1)+u5*t5**r5+w5)**2 + (w5+x5+1)**
2+(t5**r5)**2 + (u5+y5+1)**2+(t5)**2 + (u5)**2+(2*z5+1)**2 + (
r6)**2+(L6)**2 + (s6)**2+(i)**2 + (t6)**2+(2**s6)**2 + ((1+t6)
```

L'aritmetizzazione di Ω può essere ottenuta sostituendo la rappresentazione binaria di un programma specifico per il calcolo della k-esima cifra della successione Ω_n a una variabile nell'equazione derivata da un programma per macchine a registro con un «compilatore». Ω_n costituisce un'approssimazione di «n-esimo ordine» a Ω, la probabilità che un calcolatore si fermi qualora i bit che ne costituiscono il programma siano stati determinati attraverso una procedura di tipo casuale, per esempio lanciando una moneta.

Ci sono altri problemi, in altri campi della scienza, che possono trarre beneficio da queste concezioni sui fondamenti della matematica? Ritengo che la teoria algoritmica dell'informazione possa interessare la biologia. I geni che regolano lo sviluppo di un embrione sono di fatto un programma per calcolatore per la costruzione di un organismo. Si può concepire una misurazione della «complessità» di questo programma per calcolatore biochimico in termini analoghi a quelli elaborati per quantificare il contenuto informativo di Ω.

Anche se Ω è del tutto casuale (o infinitamente complesso) e non può essere calcolato esattamente, avendo a disposizione una quantità infinita di tempo si può arrivare a un'approssimazione del suo valore con la precisione che si vuole. Penso che in modo analogo si possa arrivare a un'approssimazione della complessità degli organismi viventi. L'avvicinamento a Ω di una successione di Ω_n può essere visto come una metafora dell'evoluzione della complessità biologica.

Un tempo, J. von Neumann sfidò i ricercatori a trovare una teoria matematica astratta relativa all'origine e all'evoluzione della vita. La teoria algoritmica dell'informazione può forse indicare come si possa operare per progredire in tal senso.

BIBLIOGRAFIA

CHAITIN GREGORY J., *Algorithmic Information Theory*, Cambridge University Press, 1987.

CHAITIN GREGORY J., *Information, Randomness & Incompleteness*, World Scientific Publishing Co. Pte. Ltd., 1987.

STEWART IAN, *The Ultimate in Undecidability* in «Nature», 232, n. 6160, 10 marzo 1988.

Diavoletti, macchine e il secondo principio

Lo studio dei limiti teorici dei calcolatori ha inaspettatamente fornito una soluzione all'enigma del diavoletto ipotizzato da Maxwell nel 1871 come una possibile violazione del secondo principio della termodinamica

di Charles H. Bennett

Un aspetto del secondo principio della termodinamica è che tutte le macchine che producono differenze di temperatura, come i frigoriferi, richiedono energia per funzionare; per converso è possibile utilizzare una differenza di temperatura per produrre lavoro, come fa, per esempio, una macchina a vapore, che sfrutta la differenza di temperatura fra la caldaia calda e il condensatore freddo. Eppure, nel 1871, il fisico scozzese James Clerk Maxwell, nel suo libro *Teoria del calore*, ipotizzò che una creatura abbastanza piccola da poter vedere e maneggiare singole molecole potesse sfuggire alle conseguenze di questo principio. Potrebbe infatti produrre e mantenere una differenza di temperatura senza alcuna spesa di energia:

«... se concepiamo un essere con una vista così acuta da poter seguire ogni molecola nel suo movimento, pur avendo le medesime nostre limitazioni per quanto riguarda altri attributi, questi potrebbe fare ciò che a noi oggi è impossibile. Abbiamo visto che, in un recipiente pieno d'aria a temperatura uniforme, le molecole si muovono con velocità tutt'altro che uniforme... Ora immaginiamo che tale recipiente sia diviso in due parti, *A* e *B*, mediante un setto nel quale sia praticato un forellino e che un essere, in grado di vedere le singole molecole, apra e chiuda tale foro, in modo da consentire il passaggio da *A* a *B* solo alle molecole più veloci e il passaggio da *B* ad *A* solo a quelle più lente. Egli riuscirà quindi, senza compiere alcun lavoro, ad alzare la temperatura di *B* e ad abbassare quella di *A*, in contraddizione con il secondo principio della termodinamica.»

Questo «essere» è stato presto chiamato il diavoletto di Maxwell per la sua capacità di sovvertire l'ordine naturale delle cose. Il risultato principale della sua attività sarebbe quello di abolire la necessità delle fonti di energia più o meno tradizionali, come il petrolio, l'uranio e la luce solare. Macchine di tutti i tipi potrebbero funzionare senza batterie, serbatoi di combustibile o linee elettriche. Per esempio, il diavoletto potrebbe consentire il funzionamento di una macchina a vapore senza che consumi combustibile, mantenendo indefinitamente la caldaia calda e il condensatore freddo.

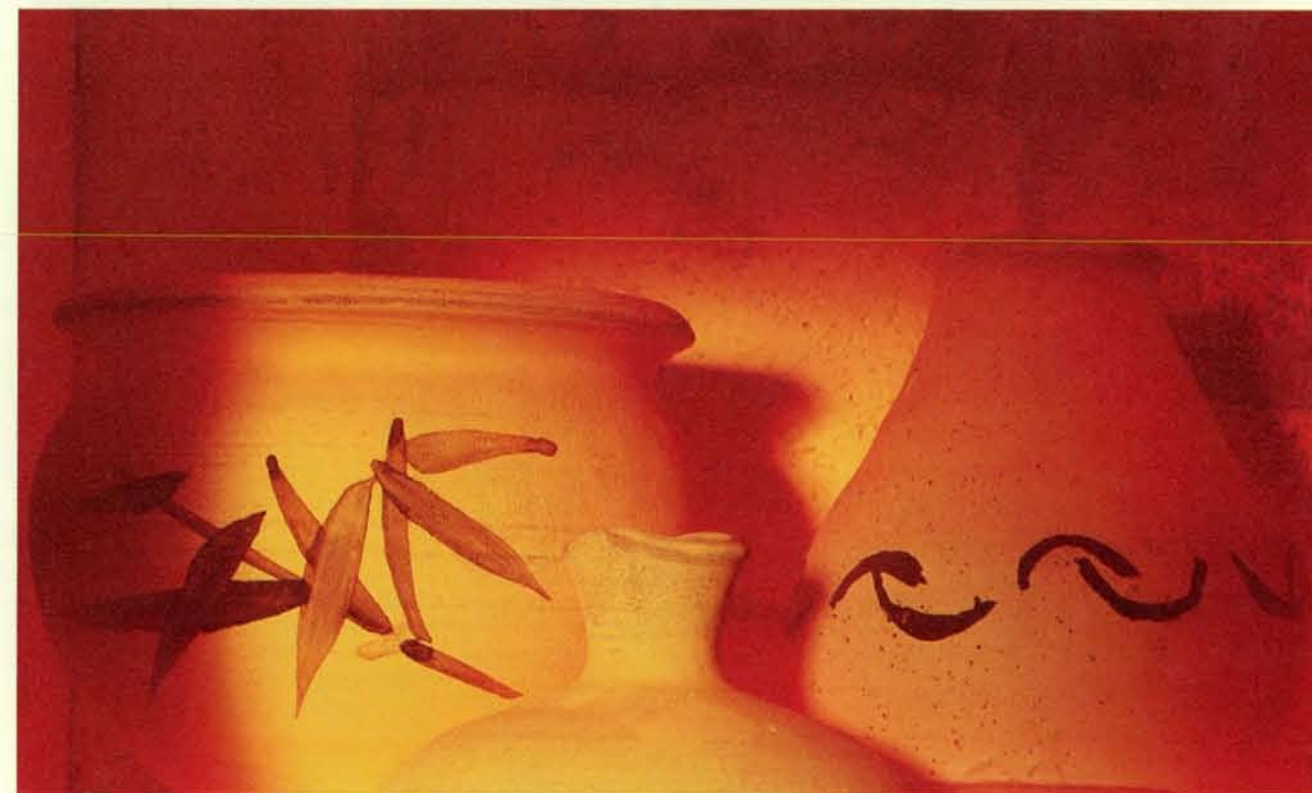
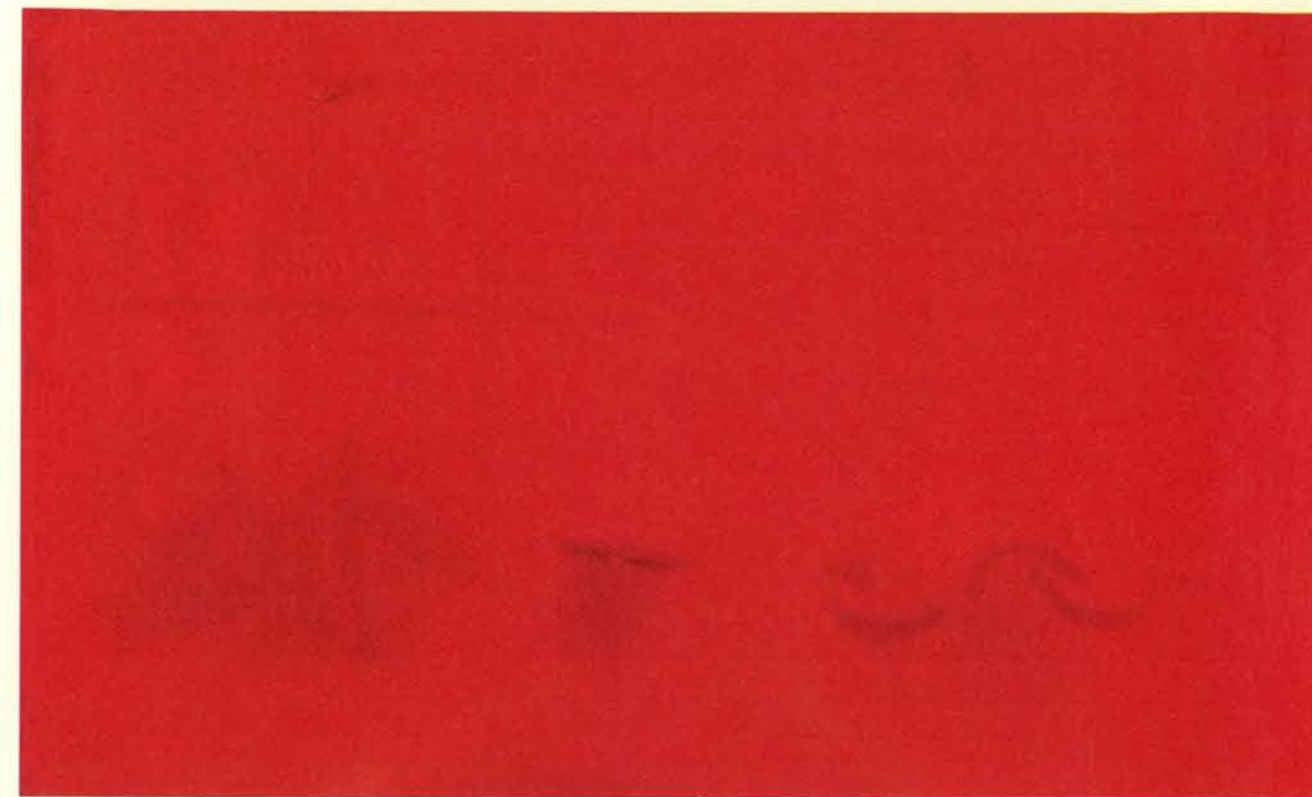
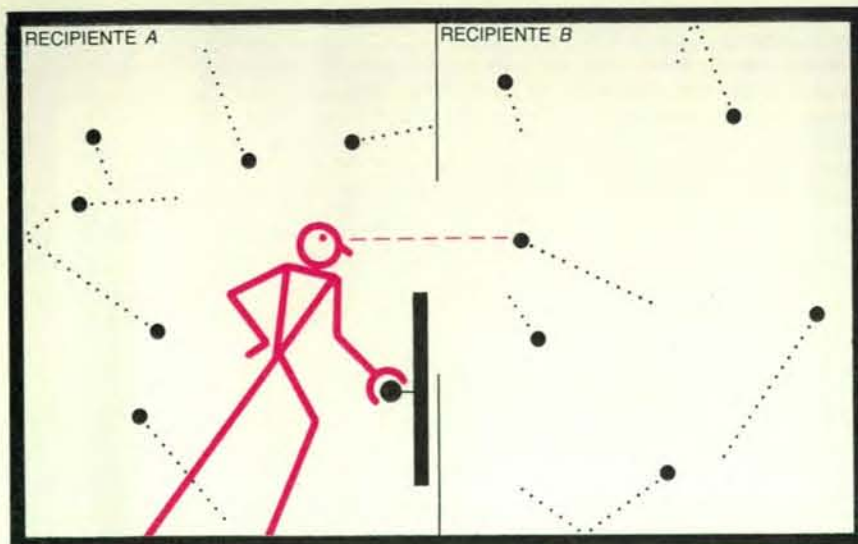
Per salvare il secondo principio sono stati indicati diversi motivi per i quali il diavoletto non potrebbe agire come ipotizzato da Maxwell, ma, sorprendentemente, tutte queste proposte non reggono a una critica approfondita. Spesso gli errori sono dovuti a un uso non corretto dei progressi fatti in altri settori della fisica. Molti hanno infatti ritenuto erroneamente che il diavoletto di Maxwell non possa operare come previsto a causa di alcune limitazioni imposte dalla meccanica quantistica.

Ma la vera ragione per cui il diavoletto di Maxwell non può violare il secondo principio è stata scoperta solo recentemente. Si tratta del risultato inatteso di un filone di ricerca completamente diverso, indirizzato alla valutazione dell'energia necessaria per il funzionamento dei calcolatori elettronici.

Fin dai tempi di Maxwell sono state proposte numerose versioni del diavoletto termodinamico. In una delle più semplici, esso produce una differenza di pressione, anziché di temperatura, consentendo a tutte le molecole, indipendentemente dalla loro velocità, di passare da *B* ad *A*, ma impedendone il passaggio nel verso opposto. Dopo un po' la maggior parte delle molecole si concentreranno in *A*, mentre in *B* si produrrà un vuoto parziale. Questo diavoletto appare molto più verosimile della creatura originale di Maxwell, dato che non è necessario che sia in grado di vedere e di pensare. Non vi è un motivo immediatamente evidente che impedisca di realizzare questo diavoletto - una valvola a flusso unidirezionale per le molecole - utilizzando dispositivi inanimati, come un minuscolo battente a molla.

Come il diavoletto di Maxwell, questo dispositivo «a pressione» potrebbe costituire una sorgente di illimitata energia per molte macchine. Per esempio, i martelli pneumatici usati per i lavori stradali vengono azionati con aria compressa da un motocompressore che funziona a benzina. Una valvola a flusso unidirezionale consentirebbe l'eliminazione del compressore, raccogliendo nel serbatoio a pressione l'aria circostante, senza alcun consumo di energia.

Si potrebbe pensare che un simile dispositivo violi il principio della conservazione dell'energia, o primo principio della termodinamica, ma ciò non è vero. L'energia necessaria per perforare il cemento potrebbe essere presa dall'energia interna dell'aria raccolta dalla valvola unidirezionale. La temperatura dell'aria, infatti, diminuirebbe attraversando il martello pneumatico. Il primo principio della termodinamica non vieta a una macchina di soddisfare le proprie esigenze energetiche estraendo calore dall'ambiente circostante o, addirittura, utilizzando il calore prodotto per attrito o recuperabile dai propri scarichi. È il secondo principio che nega la possibilità di realizzare simili macchine.

Per studiare da vicino il meccanismo d'azione del diavoletto occorre quindi comprendere alcune sottigliezze del secondo principio della termodinamica. Questo principio venne espresso inizialmente come una limitazione delle possibili trasformazioni del calore e del lavoro, ma oggi viene considerato essenzialmente come un'affermazione relativa all'aumento di disordine nell'universo. In base al secondo principio, l'entropia, ov-

La luminosità uniforme dell'interno di una fornace (*in alto*) è la dimostrazione di una conseguenza del secondo principio della termodinamica: è impossibile distinguere gli oggetti in un ambiente a temperatura uniforme senza l'ausilio di una sorgente di luce esterna di temperatura superiore a quella dell'ambiente. All'interno della fornace gli oggetti emettono luce in maniera tale che le loro superfici, indipendentemente dalla riflettanza e dal colore di ognuna di esse, appaiono ugualmente luminose. Infatti, se un oggetto apparisse più scuro di quelli circostanti, assorbirebbe energia a spese dei suoi vicini; aumenterebbe quindi la sua temperatura mentre gli oggetti intorno a esso si raffredderebbero. Ma il secondo principio della termodinamica afferma che oggetti a temperature inizialmente uguali non possono assumere spontaneamente temperature diverse. Nell'illustrazione in alto si osserva un po' di contrasto residuo perché la temperatura della fornace non è perfettamente uniforme. Le differenze intrinseche di riflettanza diventano evidenti quando gli oggetti vengono illuminati da una sorgente esterna (*in basso*).

Il diavoletto di Maxwell, descritto nel 1871 da James Clerk Maxwell, sembra in grado di violare il secondo principio della termodinamica. Il diavoletto aziona una porticina su un setto posto fra due recipienti che contengono gas alla stessa temperatura e pressione. Egli osserva le molecole che si avvicinano al foro e apre o chiude la porticina lasciando passare le molecole più veloci dal recipiente *A* a quello *B*, ma non viceversa, e quelle più lente solo da *B* ad *A*. Il recipiente *B* si riscalda, mentre l'altro si raffredda. Per il secondo principio occorre lavoro per produrre una differenza di temperatura, ma il lavoro per azionare la porta può essere reso piccolo a piacere.

vero il disordine dell'universo nel suo complesso, non può diminuire. Ne segue quindi che sono possibili solo due tipi di processi: quelli nel corso dei quali l'entropia dell'universo aumenta e quelli durante i quali essa rimane costante. I primi vengono detti processi irreversibili, perché il ripercorrere a ritroso la medesima via seguita all'andata violerebbe il secondo principio. I secondi sono detti invece processi reversibili. È possibile diminuire l'entropia di un sistema compiendo su di esso del lavoro, ma così facendo si aumenta l'entropia di qualche altro sistema o quella dell'ambiente che circonda il primo in misura maggiore, o al più uguale, della diminuzione ottenuta.

Un classico processo irreversibile, che aiuta a definire con maggior precisione il concetto di entropia, è noto come espansione libera. Supponiamo che un recipiente pieno di gas sia separato con un setto da un recipiente vuoto delle medesime dimensioni. Se si pratica un forellino nel setto il gas sfugge, cioè si espande liberamente, distribuendosi equamente fra i due recipienti.

Il motivo per il quale le molecole del gas si diffondono fino a riempire entrambi i recipienti ha, se è possibile tale distinzione, un carattere più matematico che fisico. Il numero delle molecole dai due lati del setto tende a divenire uguale non perché le molecole si respingano reciprocamente e quindi tendano a distanziarsi il più possibile, ma piuttosto perché le collisioni fra di loro e con le pareti dei recipienti ne provocano la distribuzione casuale in tutto lo spazio disponibile, fino a che metà si trovi da un lato del setto e metà dal lato opposto.

Dato che la diffusione delle molecole è dovuta più al caso che alla repulsione, esiste una possibilità che tutte le molecole ritornino spontaneamente nel recipiente dal quale sono venute. Ma la probabilità che ciò accada con n molecole equivale alla probabilità di ottenere tutte «teste» nel lancio di n monete: $1/2^n$. Quindi, per ogni numero di molecole piuttosto grande (e ve ne sono circa 300 000 000 000 000 000 000 000 in un grammo di idrogeno), l'espansione libera risulta effettivamente un processo irreversibile, ossia un processo la cui reversibilità, per quanto possibile, è talmente improbabile da consentire di affermare con certezza che non verrà mai osservata.

Lo stato disordinato, quello nel quale il gas è diffuso in entrambi i recipienti invece di rimanere raccolto in uno solo di essi, è più probabile dello stato ordinato. Ciò significa che le configurazioni in cui le molecole occupano entrambi i recipienti sono più numerose di quelle in cui ne occupano uno solo, esattamente come, quando si lanciano 100 monete, vi sono più modi di ottenere 50 volte testa e 50 volte croce di quanti non ve ne siano di ottenere 100 volte testa e mai croce. Affermando che l'entropia dell'universo tende ad aumentare, il secondo principio si limita a dire che, con il trascorrere del tempo, l'universo tende a disporsi nello stato più probabile.

È possibile quantificare quest'idea? In altre parole, è possibile dire di quanto è aumentato il disordine del gas in seguito alla sua diffusione in entrambi i recipienti? Consideriamo una singola molecola del gas. Una molecola che può vagare in entrambi i recipienti è in grado di occupare un numero di posizioni doppio di quello possibile per una molecola confinata in un solo recipiente. Se nei due recipienti vi sono due molecole, ognuna di esse ha a disposizione il doppio delle configurazioni possibili per una molecola in un solo recipiente e quindi il sistema ha il quadruplo (2 × 2) delle configurazioni possibili. Se vi sono tre molecole il sistema ha otto volte (2 × 2 × 2) il numero di configurazioni possibili in un solo recipiente.

Generalizzando, se vi sono n molecole di gas in due recipienti, queste hanno a disposizione 2^n volte il numero di configurazioni possibili per n molecole dello stesso gas in un singolo recipiente. Si dice che il gas nei due recipienti ha un numero di «stati accessibili» pari a 2^n volte quanti ne abbia in un singolo recipiente. Nella maggior parte dei sistemi il numero degli stati accessibili dipende esponenzialmente dal numero delle molecole presenti.

L'entropia di un sistema viene pertanto definita come il logaritmo del numero degli stati accessibili; nell'esempio del gas che si distribuisce in due recipienti, un aumento di 2^n volte del numero degli stati accessibili corrisponde a un aumento di entropia di n bit, o unità binarie. (La base del logaritmo - e, quindi, l'unità di entropia - è arbitraria. Si è convenuto di usare la base due e le unità binarie.) La scala logaritmica presenta il vantaggio di rendere l'entropia di una certa quantità di materia, come la sua massa e la sua energia, grosso modo proporzionale al numero delle molecole. Si può fare un'analogia con la memoria di un calcolatore: una memoria di n bit, a parità di altri fattori, ha dimensioni, peso e costo grosso modo proporzionali a n, sebbene il numero dei suoi stati possibili sia 2^n.

I primi enunciati del secondo principio non facevano riferimento alla casualità e al disordine; si riferivano invece al calore, al lavoro e alla temperatura. Come è possibile mettere questi concetti in relazione con la nostra definizione quantitativa di entropia?

Le molecole sono sempre in movimento. La velocità e la direzione del moto di ciascuna di esse sono casuali, ma la loro velocità media è proporzionale alla radice quadrata della temperatura (misurata a partire dallo zero assoluto). All'aumentare della temperatura, e quindi della velocità media, le velocità delle singole molecole si distribuiscono in un intervallo di valori più ampio di quanto avvenga quando la velocità media è bassa.

Quando la velocità media è elevata, ogni molecola ha a disposizione una maggiore gamma di velocità, proprio come una molecola nel dispositivo formato da due recipienti può occupare un numero più elevato di posizioni che in un solo recipiente. Ne segue che a temperature elevate vi è un maggior numero di stati accessibili di quanti ne siano disponibili a temperature inferiori. A temperature elevate il movimento è più disordinato perché è più difficile prevedere quale sarà la velocità di ciascuna molecola.

Il disordine nel movimento e nelle posizioni delle molecole contribuiscono entrambi a determinare l'entropia di un sistema. L'entropia di un gas può aumentare sia in conseguenza di un aumento del volume occupato (espansione), sia in seguito a un aumento di temperatura: entrambi producono un movimento molecolare più disordinato.

Un flusso di calore trasporta quindi entropia. Più precisamente ne trasporta una quantità proporzionale al flusso di calore diviso per la temperatura del flusso stesso. Quindi un passaggio di calore da un corpo caldo a un corpo freddo aumenta l'entropia di quest'ultimo più di quanto non diminuisca l'entropia del primo: la quantità di calore (flusso) che esce dal corpo caldo ed entra in quello freddo è la stessa, ma nel calcolo della diminuzione di entropia del corpo caldo si divide per una temperatura più alta di quella usata nel calcolo dell'aumento di entropia del corpo freddo. Un generico flusso di calore da un corpo caldo a uno più freddo aumenta quindi l'entropia dell'universo nel suo complesso.

Questa definizione di entropia permette di comprendere meglio perché il diavoletto di Maxwell violi il secondo principio. Con la sua capacità di selezionare le molecole, esso produce un flusso di calore dal recipiente A al recipiente B anche quando il secondo è divenuto più caldo del primo; provoca perciò una diminuzione dell'entropia di A maggiore dell'aumento di entropia di B e, di conseguenza, una diminuzione dell'entropia dell'universo, cosa impossibile dal punto di vista termodinamico.

Descrivendo il suo diavoletto, Maxwell affermò chiaramente di essere convinto della validità del secondo principio. Egli suggerì che gli esseri umani non possono violarlo semplicemente perché non sono in grado di vedere e maneggiare singole molecole. Questo non è un «esorcismo» del tutto soddisfacente perché lascia aperta la possibilità che un essere avente le capacità del diavoletto possa violare il secondo principio.

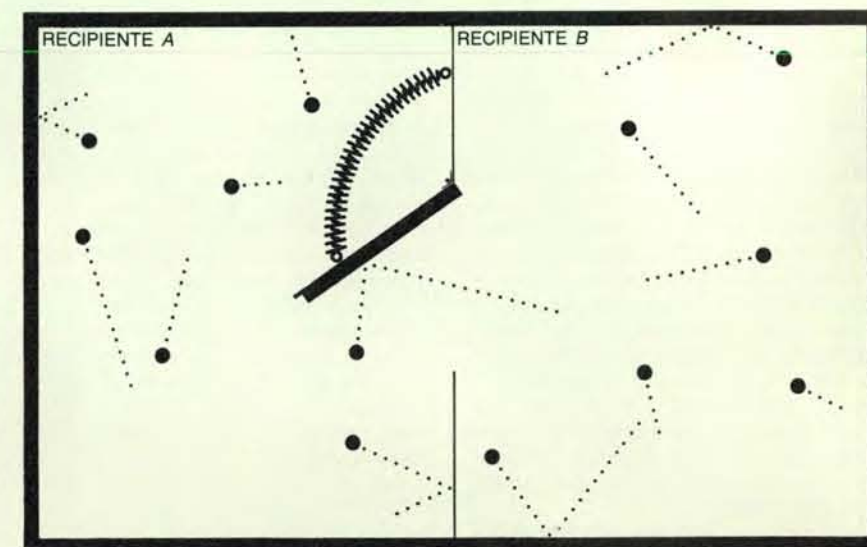
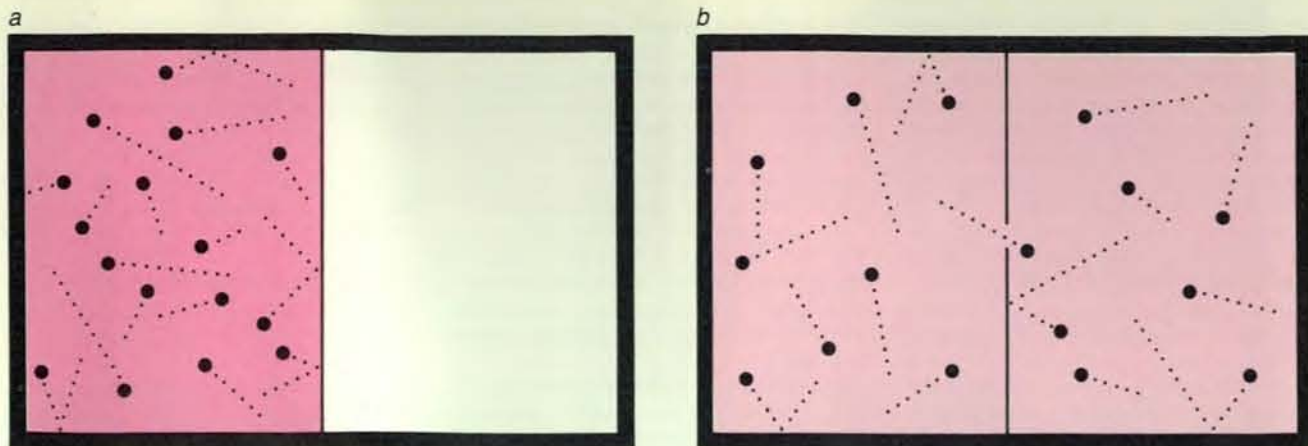
Un metodo per scoprire i motivi per i quali il diavoletto non può esistere sta nell'analizzare e confutare semplici marchingegni inanimati che potrebbero simulare l'attività del diavoletto, come la porticina a molla citata in precedenza, che funziona come una valvola a flusso unidirezionale.

Immaginiamo che la porta si apra verso sinistra. Se questa funziona come si suppone faccia il diavoletto, ogni volta che una molecola del recipiente di destra urta la porta, questa si apre e la molecola passa nel recipiente di sinistra. Al contrario, quando la porta viene urtata da una molecola del recipiente di sinistra, si serra contro il battente e non consente il passaggio. Alla fine tutte le molecole sono intrappolate a sinistra e il diavoletto ha quindi compresso il gas (diminuendone l'entropia) senza compiere lavoro.

Qual è l'errore? Innanzitutto la molla che trattiene la porta deve essere piuttosto debole. Il lavoro necessario per aprire la porta opponendosi alla forza della molla deve essere paragonabile all'energia cinetica media di una molecola del gas. Nel 1912 Marian Smoluchowski notò che la porta, ripetutamente urtata dalle molecole, finisce per acquistare un'energia cinetica (sotto forma di energia termica) conseguente al suo movimento casuale. Questa energia sarà del medesimo ordine di grandezza dell'energia di agitazione termica delle molecole del gas e quindi la porta si aprirà e chiuderà casualmente (ricordiamo che la porta è molto piccola), sbattendo alternativamente contro lo stipite e spalancandosi in opposizione alla forza della molla.

Quando la porta è aperta non può ovviamente fungere da valvola, dato che le molecole possono passare liberamente in entrambe le direzioni. Si potrebbe ancora sperare che la porta funzioni come previsto, sia pure con scarsa efficienza, intrappolando almeno una piccola quantità di gas nel recipiente di sinistra, ma in realtà non può fare neppure questo. Ogni tendenza della porta a funzionare come valvola a flusso unidirezionale, aprendosi per lasciare passare le molecole da destra a sinistra, è esattamente bilanciata dalla tendenza contraria a chiudersi davanti a una molecola, spingendola da sinistra a destra, aiutata in questo dalla forza della molla.

I due processi, il passaggio di una molecola dal recipiente di destra verso quello di sinistra e la porta che la spinge da sinistra a destra, sono l'uno il contrario dell'altro: la ripresa cinematografica di uno dei due, proiettata all'indietro, apparirebbe come una ripresa dell'altro. In un ambiente a temperatura e pressione costanti, i due processi si verificherebbero con la stessa frequenza e la possibilità che la porta funzioni da valvola a flusso unidirezionale sarebbe nulla.

Ovviamente apparecchiature di questo tipo funzionano perfettamente in ambienti nei quali la pressione sui due lati della porta è diversa. Versioni ingrandite della porticina a molla si possono vedere nei ventilatori impiegati per espellere l'aria viziata dai locali pubblici senza consentire il passaggio di correnti di aria dall'esterno quando il ventilatore è spento. Una versione microscopica funzionerebbe in maniera simile, permettendo alle molecole di passare quando vi è un eccesso di pressione su un lato della porta, ma chiudendo il flusso quando la pressione è maggiore dall'altro lato. Questi dispositivi non violano il secondo principio perché non fanno altro che permettere a pressioni diverse di uguagliarsi; essi non potrebbero mai generare una differenza di pressione.

Con una porticina a molla si può realizzare una versione automatica del diavoletto di Maxwell che produce una differenza di pressione anziché di temperatura. Fra due recipienti contenenti inizialmente gas alla stessa pressione e temperatura vi è un foro munito di una porticina a molla. La porticina si apre in una sola direzione, per lasciare passare le molecole dal recipiente *B* al recipiente *A*, ma non viceversa. Si può pensare che, alla fine, le molecole si accumulino in *A* a spese di *B*, producendo una differenza di pressione. Ma, in pratica, questo non può avvenire. La porticina, riscaldata dagli urti con le molecole, prende ad aprirsi e a chiudersi casualmente per agitazione termica. Quando è aperta non può fungere da valvola a flusso unidirezionale e quando si chiude può spingere una molecola da *A* a *B*. Il secondo processo avviene altrettanto frequentemente di quello nel quale una molecola di *B* spinge la porta per passare in *A*.

L'espansione libera di un gas è un processo termodinamicamente irreversibile durante il quale l'entropia, ovvero il disordine, dell'universo cresce. Un gas è inizialmente confinato in uno di due recipienti (*a*) dell'apparato sperimentale illustrato. Quando il setto fra i recipienti viene forato, le molecole fluiscono da un recipiente all'altro fino a quando il loro numero da entrambi i lati del setto è circa uguale (*b*).

Se un semplice analogo meccanico del diavoletto non può funzionare, forse la cosa è possibile per un essere intelligente. In effetti, qualche tempo dopo che Maxwell aveva descritto il suo diavoletto, diversi ricercatori giunsero a pensare che il fattore critico che consentiva al diavoletto di operare fosse appunto l'intelligenza. Per esempio, in un articolo del 1914, Smoluchowski scrisse: «Per quanto ne sappiamo oggi, non esiste alcuna macchina automatica in grado di realizzare un moto perpetuo, malgrado la presenza di fluttuazioni molecolari, ma un simile apparecchio potrebbe forse funzionare, qualora fosse azionato da esseri intelligenti».

Il fisico Leo Szilard affrontò una trattazione quantitativa del problema in un lavoro pubblicato nel 1929 il cui titolo può essere così tradotto: *Sulla diminuzione di entropia di un sistema termodinamico in seguito all'intervento di esseri intelligenti*. Sebbene il titolo sembri ammettere la possibilità che un diavoletto intelligente violi il secondo principio, nello scritto si confuta tale ipotesi per sostenere che nessun essere, intelligente o meno, può comportarsi come il diavoletto di Maxwell. Szilard riteneva che le osservazioni, o misurazioni, che il diavoletto deve compiere (per esempio, per determinare da che parte arriva una molecola) non possano venire effettuate senza spendere una quantità di lavoro sufficiente a produrre un aumento di entropia tale da impedire la violazione del secondo principio.

Szilard aveva preso in considerazione un diavoletto che differiva sotto vari aspetti da quello proposto da Maxwell e che venne quindi chiamato macchina di Szilard. (La macchina che sto per descrivere differisce però leggermente da quella di Szilard.) Il componente principale della macchina è un cilindro nel quale è contenuta una sola molecola, soggetta ad agitazione termica casuale. Ciascuna estremità del cilindro è chiusa da un pistone e a metà del cilindro può essere inserito un setto mobile sottile in grado di intrappolare la molecola in una delle due metà. La macchina è pure munita di dispositivi di osservazione per determinare in quale metà del cilindro si trovi la molecola e di un sistema di memoria per registrare l'informazione. Il ciclo della macchina consiste di sei fasi. Nella prima si inserisce il setto, intrappolando la molecola in una delle due metà del cilindro. Szilard riteneva che, in via teorica, il lavoro per l'inserzione del setto potesse essere reso trascurabile.

Nella fase successiva la macchina determina in quale delle due metà del cilindro si trovi la molecola. Il sistema di memoria può trovarsi in tre stati distinti. Uno stato indifferente mostra che non è stata compiuta alcuna osservazione, uno stato *S* indica che la molecola è intrappolata a sinistra e uno stato *D* corrisponde alla situazione in cui la molecola è a destra. Quando si effettua la determinazione, la memoria passa dallo stato indifferente a uno degli altri due.

La terza fase del ciclo, che può essere considerata una compressione, dipende dalle conoscenze acquisite nella fase precedente. Il pistone posto nella metà nella quale non si trova la molecola viene spinto fino a toccare il setto. Diversamente dalla fase di compressione di un motore a combustione interna, questa compressione non richiede lavoro, perché il pistone «comprime» uno spazio vuoto. La molecola, intrappolata dall'altra parte del setto, non può opporre resistenza al movimento del pistone.

Successivamente, nella quarta fase, viene rimosso il setto, consentendo alla molecola di urtare il pistone che è appena avanzato e di esercitare su di esso una pressione.

Nella quinta fase, che può venire detta fase di potenza, la pressione della molecola riporta il pistone nella posizione originale, compiendo lavoro su di esso. L'energia che la molecola fornisce al pistone viene rimpiazzata dal calore diffuso dall'ambiente attraverso la parete del cilindro, perciò la molecola continua a muoversi alla medesima velocità media. Il risultato è quindi quello di convertire il calore sottratto all'ambiente in lavoro meccanico eseguito sul pistone.

Nella sesta e ultima fase la macchina azzera la sua memoria, riportandola nello stato indifferente. Essa ha ora assunto la stessa configurazione che aveva all'inizio e il ciclo può venire ripetuto.

Complessivamente le sei fasi sembrano essere riuscite a convertire il calore sottratto all'ambiente in lavoro, facendo tornare il gas e la macchina esattamente allo stato iniziale. Se non è avvenuto alcun altro mutamento, nel corso del ciclo compiuto dalla macchina l'entropia dell'universo è diminuita. In linea di principio, il ciclo potrebbe venire ripetuto quante volte si vuole, ottenendo come risultato una violazione arbitrariamente grande del secondo principio.

La soluzione del paradosso proposta da Szilard consisteva nel postulare che l'atto della misurazione, nel corso del quale viene determinata la posizione della molecola, implichi un aumento di entropia tale da compensare la diminuzione di entropia ottenuta durante la fase di potenza. Szilard rimase alquanto vago a proposito della natura e della collocazione dell'aumento di entropia ma, negli anni che seguirono la pubblicazione del suo lavoro, parecchi fisici, e fra questi in particolare Léon Brillouin (che nel 1956 scrisse un volume sulla scienza e la teoria dell'informazione tradotto nel 1962 in inglese) e Denis Gabor (l'inventore dell'olografia), cercarono di dare delle basi alla postulata irreversibilità della misurazione. In particolare, tentarono di valutare il costo, in termini di energia e di entropia, dell'osservazione di una molecola, colpendola con un fascio di luce e osservandone la riflessione.

Brillouin e Gabor basarono il loro lavoro su una teoria che si era sviluppata fin dai tempi di Maxwell: la teoria quantistica della radiazione. Secondo la teoria ondulatoria classica della radiazione (alla quale Maxwell diede contributi fondamentali) l'energia di un raggio di luce può essere arbitrariamente piccola. Ma, secondo la teoria quantistica, la luce è costituita da pacchetti di energia, i fotoni. L'energia di un fotone dipende dalla sua lunghezza d'onda (legata alla sensazione di colore che esso produce) ed è impossibile osservare meno di un singolo fotone. Le argomentazioni di Brillouin si basavano sul fatto che per osservare una molecola questa deve diffondere almeno un fotone del fascio di luce usato per osservarla e quando l'energia del fotone viene dissipata in calore si ha un aumento di entropia almeno uguale alla diminuzione ottenibile dalla macchina di Szilard in seguito all'informazione acquisita sulla posizione della molecola che ha interagito con la luce.

Perché allora non usare per l'osservazione un fascio di fotoni di energia molto bassa? L'idea non funziona per un'altra conseguenza, più complessa, della teoria quantistica. Questa prevede che un recipiente le cui pareti e il cui interno siano a temperatura uniforme si riempia di un «gas» di fotoni. Le lunghezze d'onda dei fotoni dipendono dalla temperatura del recipiente. Questo gas di fotoni è responsabile della luminosità uniforme di colore giallo o arancione che si osserva all'interno di una fornace. (A temperatura ambiente i fotoni hanno per lo più energie corrispondenti alla zona infrarossa dello spettro e quindi non sono visibili.)

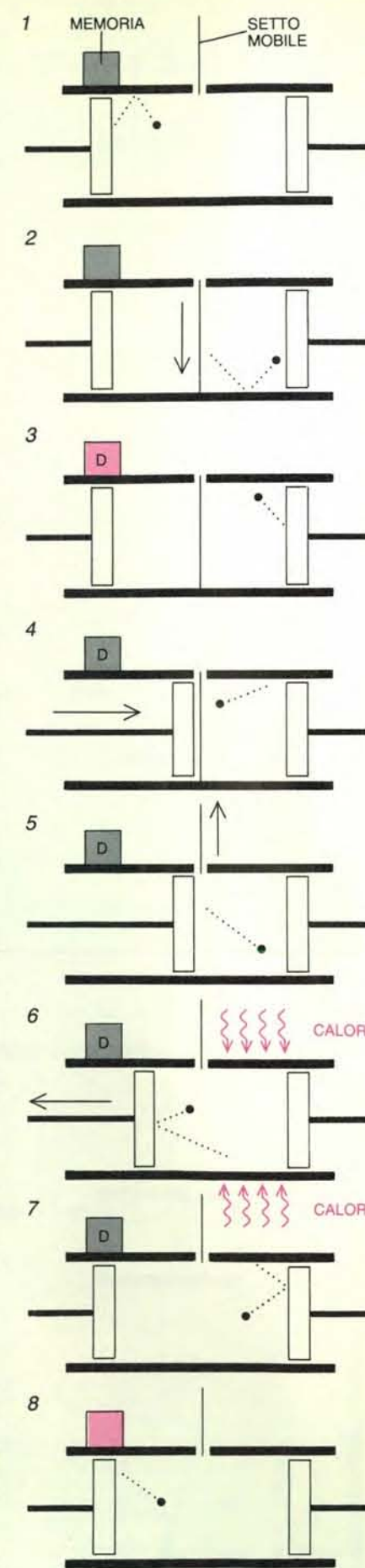
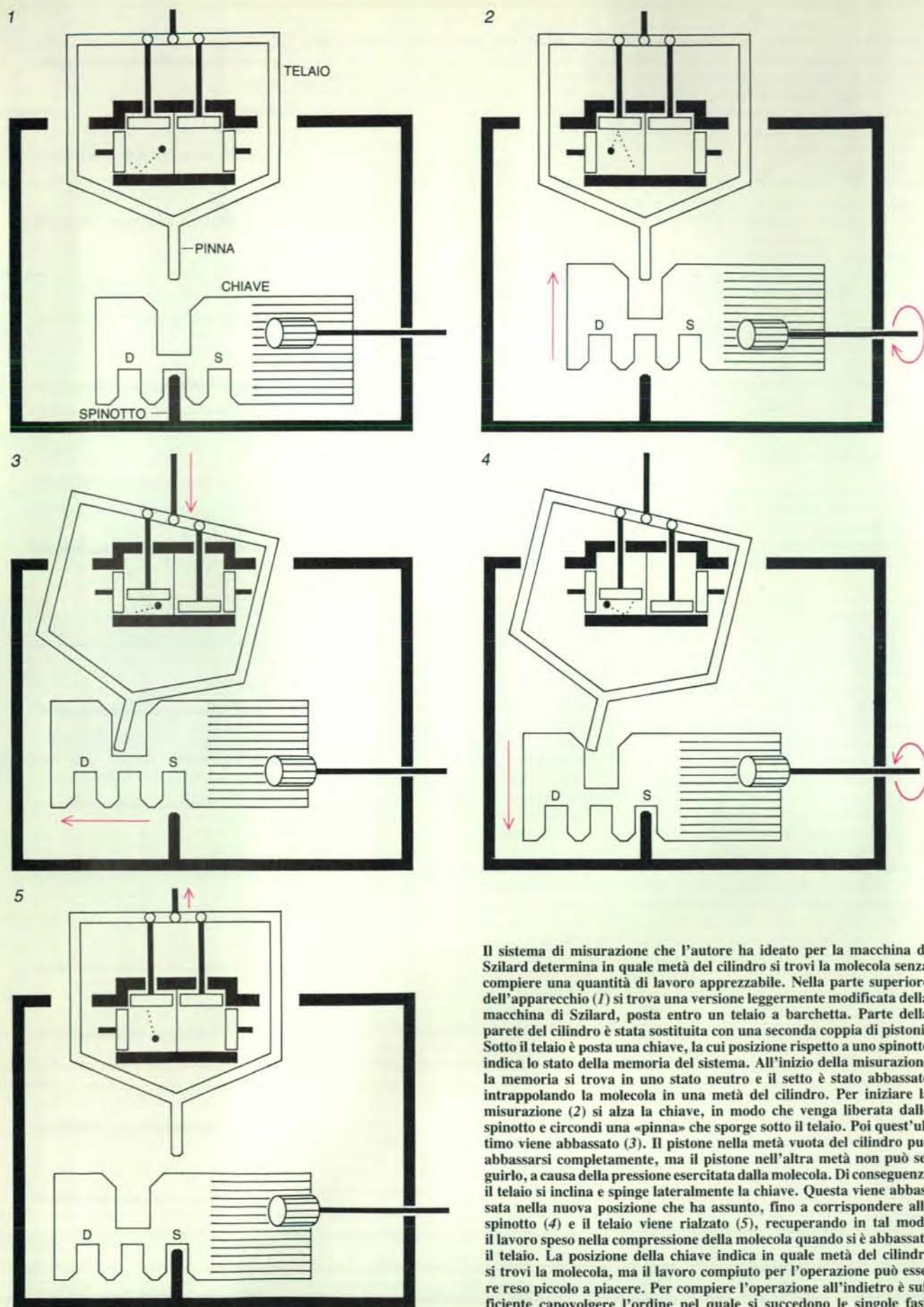
A prima vista, il gas di fotoni sembrerebbe un'ottima sorgente di luce per mezzo della quale il diavoletto potrebbe osservare le molecole, risparmiando così il costo entropico di una lampada. Ma una conseguenza sorprendente del secondo principio, scoperta nel 1859 da Gustav Robert Kirchhoff, è che risulta impossibile vedere gli oggetti contenuti in un recipiente a temperatura uniforme usando la radiazione luminosa emessa dal recipiente stesso. In effetti, se si osserva, per esempio, l'interno di un forno nel quale vengono cotti alcuni vasi, si vede una luminescenza arancione uniforme, quasi completamente priva di contrasto, anche se i vasi hanno colori e superfici completamente diversi.

Gli oggetti nel forno appaiono come se avessero tutti la medesima luminosità e tonalità di colore, ma si può verificare che non è così illuminandoli dall'esterno con un fascio di luce sufficientemente intenso. La ragione per la quale gli oggetti nel forno quasi scompaiono è che quelli scuri (che riflettono meno la luce) emettono una luce più intensa di quelli chiari (più riflettenti): in questo modo, l'intensità totale della luce (emessa e riflessa) che proviene da ogni oggetto è sempre la stessa.

Per comprendere perché debba verificarsi questo strano livellamento, supponiamo che esso non avvenga e pensiamo a quali potrebbero essere le conseguenze per il secondo principio. Immaginiamo che due oggetti, per esempio un vaso e una pentola, siano posti uno vicino all'altro in un forno a temperatura uniforme. Se l'intensità della luce che lascia il vaso in direzione della pentola fosse maggiore di quella cha va dalla pentola al vaso, si avrebbe un flusso di energia dal vaso alla pentola. Questa si scalderebbe mentre il primo si raffredderebbe.

Quindi, senza spendere lavoro, due oggetti inizialmente alla stessa temperatura assumerebbero temperature diverse, in contraddizione con il secondo principio, proprio come se un diavoletto di Maxwell si fosse interposto fra di essi. Allora, se il secondo principio deve essere valido, gli oggetti in un forno a temperatura uniforme non possono avere luminosità superficiali diverse.

Per vedere gli oggetti contenuti in una fornace è quindi indispensabile illuminarli dall'esterno, per esempio con una lampada il cui filamento sia a temperatura superiore di quella esistente dentro il forno. Nella nostra esperienza quotidiana una simile sorgente di luce, il Sole, ci consente di vedere gli oggetti entro contenitori che si trovano tutti a temperatura ambiente.

Brillouin, Gabor e altri, armati della loro conoscenza delle proprietà del gas di fotoni, sostenevano che il diavoletto di Maxwell non potesse vedere le molecole da selezionare senza impiegare una qualche sorgente di luce. Pertanto concludevano che esso non poteva violare il secondo principio. Ogni qual volta il diavoletto osserva una molecola, deve dissipare almeno l'energia di un fotone e questa energia deve essere superiore a

La macchina di Szilard, descritta da Leo Szilard nel 1929, sembra poter convertire il calore dell'ambiente in lavoro, violando il secondo principio. La macchina (*1*) è costituita da un cilindro con le estremità chiuse da pistoni; è munita di un setto mobile e di apparecchiature per osservare il contenuto del cilindro e memorizzare le osservazioni. Il cilindro contiene una sola molecola. All'inizio del ciclo (*2*) si abbassa il setto, intrappolando la molecola in una metà del cilindro. Il sistema di osservazione determina e memorizza la posizione della molecola (*3*) e il pistone dalla parte opposta viene spinto fino a toccare il setto (*4*). Il pistone viene spostato senza compiere lavoro, dato che si muove nel vuoto. Poi il setto viene ritirato (*5*) e la molecola urta il pistone (il gas monomolecolare si «espande»), spingendolo indietro (*6*). L'energia spesa dalla molecola nel compiere lavoro sul pistone è compensata dal calore sottratto all'ambiente. Quando il pistone è tornato nella posizione originale (*7*), la memoria viene cancellata (*8*) e il ciclo può ricominciare.

Il sistema di misurazione che l'autore ha ideato per la macchina di Szilard determina in quale metà del cilindro si trovi la molecola senza compiere una quantità di lavoro apprezzabile. Nella parte superiore dell'apparecchio (*1*) si trova una versione leggermente modificata della macchina di Szilard, posta entro un telaio a barchetta. Parte della parete del cilindro è stata sostituita con una seconda coppia di pistoni. Sotto il telaio è posta una chiave, la cui posizione rispetto a uno spinotto indica lo stato della memoria del sistema. All'inizio della misurazione la memoria si trova in uno stato neutro e il setto è stato abbassato intrappolando la molecola in una metà del cilindro. Per iniziare la misurazione (*2*) si alza la chiave, in modo che venga liberata dallo spinotto e circondi una «pinna» che sporge sotto il telaio. Poi quest'ultimo viene abbassato (*3*). Il pistone nella metà vuota del cilindro può abbassarsi completamente, ma il pistone nell'altra metà non può seguirlo, a causa della pressione esercitata dalla molecola. Di conseguenza il telaio si inclina e spinge lateralmente la chiave. Questa viene abbassata nella nuova posizione che ha assunto, fino a corrispondere allo spinotto (*4*) e il telaio viene rialzato (*5*), recuperando in tal modo il lavoro speso nella compressione della molecola quando si è abbassato il telaio. La posizione della chiave indica in quale metà del cilindro si trovi la molecola, ma il lavoro compiuto per l'operazione può essere reso piccolo a piacere. Per compiere l'operazione all'indietro è sufficiente capovolgere l'ordine nel quale si succedono le singole fasi.

un determinato valore minimo, imposto dalla temperatura del gas di fotoni entro il quale il diavoletto è immerso. Argomenti di questo genere, per quanto non rigorosi, sembravano confermare la convinzione espressa da Szilard che l'ottenimento di una determinata quantità di informazioni richiede sempre la produzione di una corrispondente quantità di entropia.

Il successivo importante passo verso l'esorcismo del diavoletto è stato un risultato collaterale delle ricerche condotte da Rolf Landauer della IBM sulla termodinamica dell'elaborazione di dati. Alcune operazioni di elaborazione, come la copiatura di dati da un sistema di memoria a un altro, sono analoghe a una misurazione, in quanto un sistema acquisisce conoscenze sullo stato di un altro. Di conseguenza, negli anni cinquanta, si conveniva che le operazioni di elaborazione fossero intrinsecamente irreversibili (nel significato termodinamico della parola), proprio come Szilard aveva sostenuto per le misurazioni. Si riteneva quindi che qualsiasi operazione su dati richiedesse la produzione e la rimozione di almeno l'equivalente di un bit di calore per ogni bit di dati elaborato. Si tratta di una quantità di calore estremamente piccola, grosso modo un decimo di miliardesimo del calore effettivamente prodotto dai circuiti elettronici esistenti.

Intorno al 1960 Landauer esaminò il problema più a fondo e stabilì che alcune operazioni hanno in effetti un costo energetico, mentre altre, fra le quali, in determinate situazioni, la copiatura di dati, sono libere da restrizioni termodinamiche fondamentali (si veda l'articolo *I limiti fisici fondamentali del calcolo* di Charles H. Bennett e Rolf Landauer in «Le Scienze» n. 205, settembre 1985).

La dimostrazione di Landauer parte dalla premessa che stati logici diversi di un calcolatore devono essere rappresentati da stati fisici diversi del suo hardware. Per esempio, ogni stato possibile della memoria di un calcolatore deve essere rappresentato da un insieme distinto di correnti, tensioni, campi e così via.

Supponiamo che un registro di memoria di n bit venga cancellato, cioè il valore di ogni sua locazione venga posto uguale a zero, indipendentemente da quello che aveva inizialmente. Prima della cancellazione, il registro nel suo insieme poteva essere in uno qualsiasi dei suoi 2^n stati possibili. Dopo l'operazione esso può essere in un solo stato. L'operazione ha quindi compresso in uno solo molti stati possibili, a somiglianza di un pistone che comprime un gas.

In base alla premessa di Landauer, per comprimere lo stato logico di un calcolatore si deve anche comprimere il suo stato fisico. Occorre quindi diminuire l'entropia dello hardware. Questa diminuzione di entropia, per il secondo principio, non può venire effettuata senza un aumento almeno identico dell'entropia dell'ambiente del calcolatore. Quindi non è possibile azzerare un registro di memoria senza produrre calore e aumentare l'entropia dell'ambiente: si tratta di un'operazione termodinamicamente irreversibile.

Landauer identificò varie altre operazioni termodinamicamente irreversibili. Tutte queste operazioni hanno in comune il fatto di scartare informazioni concernenti lo stato precedente del calcolatore. Landauer chiama tali operazioni «logicamente irreversibili».

La rilevanza di queste idee nel problema della misurazione, implicita nel lavoro di Landauer stesso e nei modelli reversibili di elaborazione sviluppati nel corso degli anni settanta da Edward Fredkin del MIT, da me e da altri, divenne chiara nel 1982, in seguito alla mia proposta di utilizzarle per dare una risposta definitiva al problema del diavoletto di Maxwell.

Consideriamo il ciclo della macchina di Szilard. L'ultima fase, durante la quale la memoria della macchina viene riportata alla condizione neutra, è logicamente irreversibile, in quanto comprime due stati possibili («la molecola è a sinistra» e «la molecola è a destra») in uno solo («la posizione della molecola non è ancora stata determinata»). Quindi la macchina non può azzerare la sua memoria senza far salire l'entropia dell'ambiente di almeno un bit. Questo fa sì che tutto il lavoro ottenuto durante la fase di potenza venga riconvertito in calore.

E la fase di misurazione? Ha anch'essa un costo energetico? In un tal caso la macchina aumenterebbe per due volte l'entropia dell'universo, una volta determinando la posizione della molecola e una volta azzerando la memoria dopo la fase di potenza. In realtà non è necessario che la misurazione sia termodinamicamente dispendiosa. Vi sono altri modi per osservare la posizione di una molecola oltre a quello di colpirla con dei fotoni. Per provare tale possibilità ho ideato un dispositivo di misura reversibile, che determina e memorizza la posizione di una molecola senza seguire un percorso irreversibile.

Abbiamo quindi stabilito perché il diavoletto non può violare il secondo principio: per osservare una molecola deve prima dimenticare i risultati delle osservazioni precedenti. Dimenticare i risultati, ovvero scartare informazioni, ha un costo energetico.

Ovviamente, se il diavoletto avesse una memoria molto grande potrebbe ricordare i risultati di tutte le sue osservazioni. Non avremmo allora alcun processo logicamente irreversibile e la macchina potrebbe convertire in lavoro, a ogni ciclo, l'equivalente di un bit di calore. Ma in questo caso il ciclo non è più un ciclo: ogni volta la memoria, inizialmente azzerata, acquisterebbe un altro bit di informazione casuale. La corretta interpretazione termodinamica di questo fatto può essere espressa dicendo che la macchina aumenta l'entropia della propria memoria per diminuire quella dell'ambiente.

L'attribuzione dell'aumento di entropia alla fase di azzeramento della memoria anziché a quella di misurazione può sembrare una formalità contabile, dato che il ciclo completo della macchina di Szilard li comprende entrambi, ma è possibile evitare parecchia confusione se si fa una distinzione chiara fra l'acquisizione di nuove informazioni e la distruzione di quelle vecchie. Non sappiamo se tale confusione esistesse nella mente di Szilard. Nella maggior parte della sua trattazione afferma che la fase irreversibile è quella della misurazione ma, a un certo punto, fa il bilancio dei cambiamenti di entropia durante il ciclo e trova, senza fare alcun commento esplicito, che l'aumento di entropia si verifica durante l'azzeramento della memoria.

Se quanti si sono occupati in seguito dell'argomento avessero proseguito nella direzione indicata da questo aspetto dello scritto di Szilard, avrebbero raggiunto le nostre stesse conclusioni di oggi. Il fatto che ciò non sia accaduto è un'ironia della storia della scienza: lo sviluppo di un settore della fisica (la teoria quantistica della radiazione) ha ritardato in modo palese il progresso in un altro settore (la termodinamica). Un aspetto della meccanica quantistica che ha rafforzato l'idea che occorra pagare un prezzo per acquisire informazioni è il principio di indeterminazione, secondo il quale alcune misure non possono venire eseguite con una precisione grande a piacere. Per quanto il principio di indeterminazione suoni simile all'ipotesi di Szilard che stabilisce un costo entropico irriducibile per le misurazioni, i due postulati sono in realtà completamente diversi. L'ipotesi di Szilard riguarda il costo energetico di una misurazione, mentre il principio di indeterminazione riguarda la stessa possibilità di eseguirla, indipendentemente dal suo costo.

Un'altra fonte di confusione è il fatto che generalmente non si pensa all'informazione come a un inconveniente. Si paga per avere un giornale, non per farselo portare via. Intuitivamente, la registrazione da parte del diavoletto degli eventi passati viene considerata come qualcosa che ha valore (o che, al massimo, è inutile). Ma il «giornale di ieri» del diavoletto (il risultato di una misurazione precedente) occupa spazio prezioso e il costo del suo svuotamento uguaglia il beneficio che il diavoletto ha ricevuto dal giornale, quando era nuovo. Forse la crescente coscienza dell'inquinamento ambientale e l'esplosione di informazioni generata dai calcolatori hanno fatto in modo che oggi paia più naturale di quanto non fosse nei primi decenni del secolo l'idea che l'informazione possa avere un «valore» negativo.

Note biografiche e bibliografiche

1. Determinismo e predicibilità

Autore

David Ruelle è docente di fisica teorica presso l'Institut des Hautes Études Scientifiques a Bures sur Yvette. Nato in Belgio, ha compiuto gli studi a Bruxelles e a Zurigo; dopo aver conseguito la laurea nel 1960 ha lavorato a più riprese all'Institute for Advanced Studies di Princeton. Ha ottenuto l'attuale incarico nel 1964. Da diversi anni si occupa in particolar modo dello studio matematico dei sistemi dinamici e dei loro rapporti con la fisica. Insieme a Floris Takens, Ruelle è stato lo scopritore delle strutture matematiche denominate «attrattori strani», che hanno assunto una considerevole importanza nella teoria della turbolenza.

Bibliografia

Ruelle David, *Les attracteurs étranges* in «La Recherche» n. 108, febbraio 1980.

Berge Pierre e Pomeau Yves, *La turbulence* in «La Recherche» n. 110, aprile 1980.

Thom René, *Halte au hasard, silence au bruit* in «Le Débat» n. 3, luglio-agosto 1980.

Prigogine Ilya e Morin Edgar, *Réponses à René Thom* in «Le Débat» n. 6, novembre 1980.

Ekeland Ivar, *Le calcul, l'imprévu*, Éditions du Seuil, 1984.

Cvitanović Predsag (a cura), *Universality in Chaos*, Adam Hilger Ldt., 1984.

2. Il caos

Autori

James P. Crutchfield, J. Doyne Farmer, Norman H. Packard e Robert S. Shaw hanno cominciato a collaborare nello studio dei sistemi caotici, fondando il Dynamical Systems Collective, mentre si stavano preparando alla laurea in fisica presso l'Università della California a Santa Cruz. Crutchfield, completati gli studi a Santa Cruz, svolge attualmente lavoro di ricerca all'Università della California a Berkeley. Farmer, diplomatosi in fisica alla Stanford University nel 1973, si è laureato a Santa Cruz nel 1981. Attualmente lavora al Los Alamos National Laboratory, dove si trova dal 1982. Packard ha conseguito la laurea a Santa Cruz nel 1982 e, prima di passare all'attuale incarico, ha svolto lavoro di ricerca presso l'Institut des Hautes Études Scientifiques e l'Institute for Advanced Studies di Princeton. Oggi lavora all'Università dell'Illinois a Urbana-Champaign nel Dipartimento di fisica e presso il centro di ricerca sui calcolatori. Shaw, diplomatosi allo Harvard College nel 1972, si è laureato nel 1980 a Santa Cruz; oltre che scienziato, è anche un ottimo pianista con predilezione per la musica classica e le improvvisazioni.

Bibliografia

Packard Norman H., Crutchfield James P., Farmer J. Doyne e Shaw Robert S., *Geometry from a Time Series* in «Physical Review Letters», 45, n. 9, 1 settembre 1980.

Abraham Ralph e Shaw Chris, *Dynamics: The Geometry of Behavior*, Aerial Press, P.O. Box 1360, Santa Cruz, California 95061, 1982-1985.

Schuster H. G., *Deterministic Chaos: An Introduction*, VCH Publishers, Inc., 1984.

Bass Thomas A., *The Eudaemonic Pie: or Why Would Anyone Play Roulette Without a Computer in His Shoe?*, Houghton-Mifflin, 1985.

Mayer-Kress G. (a cura), *Dimensions and Entropies in Chaotic Systems*, Springer-Verlag, 1986.

3. Determinismo e caos

Autore

Vincent Croquette, all'epoca della stesura di questo articolo, stava preparando la tesi di dottorato presso il Commissariat à l'Énergie Atomique de l'Orme des Merisiers, a Gif-sur-Yvette. Le sue ricerche in quel periodo riguardavano l'origine della turbolenza in un fluido sottoposto a convezione del tipo Rayleigh-Bénard. Spinto da un profondo interesse per l'esperimento, ama ideare sistemi nella pratica, sia nel campo della strumentazione sia in quello della fisica. Diplomato presso l'École Supérieure de Physique et Chimie Industrielles (ESPCI) di Parigi, Croquette è anche appassionato di vela, volo a vela e fotografia.

Bibliografia

Wilson Kenneth, *Problemi fisici con più scale di grandezza* in «Le Scienze» n. 134, ottobre 1979.

Arnold Vladimir, *Metodi matematici della meccanica classica*, Editori Riuniti, Roma, 1979.

Berge Pierre e Pomeau Yves, *La turbulence* in «La Recherche» n. 110, aprile 1980.

Arnold V. e Avez A., *Problèmes ergodiques de la mécanique classique*, Éditions Gauthier-Villars, Parigi.

Abraham R. e Marsden J., *Foundations of Mechanics*, The Benjamin Cummings Publishing Company, Inc.

4. Il moto browniano da Einstein a oggi

Autore

Bernard H. Lavenda è professore di chimica fisica presso l'Università di Camerino. Si è diplomato in chimica alla Clark University nel 1966 e laureato in chimica fisica al Weizmann Institute, in Israele, nel 1967 e all'Université Libre di Bruxelles nel 1970. Nel 1972 si è trasferito in Italia dove è stato docente di chimica statistica presso l'Università di Pisa e dal 1975 al 1980 ha insegnato in numerosi corsi di fisica presso la Facoltà di scienze dell'Università di Napoli. Il suo interesse è centrato su problematiche fondamentali della termodinamica dei processi irreversibili. È autore del libro *Termodinamica dei processi irreversibili* e coautore del testo *Introduzione alla fisica atomica e statistica*, pubblicati dall'editore Liguori di Napoli. Dal 1972 mantiene una attiva collaborazione scientifica e didattica con la società TEMA del gruppo ENI, applicando la termodinamica statistica del non-equilibrio a processi di interesse industriale.

Bibliografia

Einstein Albert, *Investigations on the Theory of Brownian Movement*, Dover Publications, 1956.

Hersh Reuben e Griego Richard J., *Moto browniano e teoria del potenziale* in «Le Scienze» n. 10, giugno 1969.

Lavenda Bernard H. e Santamato Enrico, *The Kinetic Analog of Boltzmann's Principle* in «Journal of Mathematical Physics», 23, 1981.

Santamato Enrico e Lavenda Bernard H., *The Underlying Brownian Motion of Nonrelativistic Quantum Mechanics* in «Foundations of Physics», 11, 1981.

Lavenda Bernard H. e Santamato Enrico, *Thermodynamic Criteria Governing Irreversible Processes Under the Influence of Small Thermal Fluctuations* in «Journal of Statistical Physics», 29, 1982.

Santamato Enrico e Lavenda Bernard H., *The Stochastic H-Theorem* in «Journal of Mathematical Physics», 23, novembre 1982.

Lavenda Bernard H., *The Optional Control of the Decay of Nonequilibrium Statistical Correlations* in «Journal of Mathematical Physics», 23, dicembre 1982.

5. La freccia del tempo

Autore

David Layzer è professore di astronomia alla Harvard University, dove si è diplomato in matematica nel 1947 e si è laureato in astrofisica teorica nel 1950. Dopo un periodo dedicato a studi di fisica atomica teorica si è occupato soprattutto di problemi di cosmologia. Attualmente si interessa del magnetismo solare, come pure dei problemi riguardanti l'origine dei sistemi astronomici.

Bibliografia

Gold T. e Schumacher D. L. (a cura), *The Nature of Time*, Cornell University Press, 1967.

Layzer D., *Cosmogonic Processes* in *Astrophysics and General Relativity: Vol. II* a cura di M. Chrétien, S. Deser e J. Goldstein, Gordon and Breach Science Publishers, 1970.

Landsberg P. T. (a cura), *Proceedings of the International Conference on Thermodynamics*, Cardiff (UK), 1-4 aprile 1970, in «Pure and Applied Chemistry», 22, n. 3-4, 1970.

Layzer D., *Cosmology and the Arrow of Time* in *Vistas in Astronomy: Vol. XIII*, Pergamon Press, Inc., 1972.

Whitrow G., *The Nature of Time*, Holt, Rinehart and Winston, 1973.

Davies P. C. W., *Physics of Time Asymmetry*, University of California Press, 1974.

Layzer D., *Galaxy Clustering: Its Description and Interpretation* in *Stars and Stellar Systems: Vol. IX*, University of Chicago Press, 1976.

6. Attrattori strani: enti fra ordine e caos

Autore

Douglas R. Hofstadter ha insegnato per diversi anni al Dipartimento di scienza dei calcolatori della Indiana University e attualmente è docente presso il Dipartimento di psicologia dell'Università del Michigan ad Ann Arbor. È autore di diverse opere tradotte in italiano: *L'io della mente* (Adelphi, Milano, 1985), scritto in collaborazione con il filosofo Daniel C. Dennett, *Ambigrammi* (Hopeful Monster, Firenze, 1986) e *Gödel, Escher, Bach: un'eterna ghirlanda brillante* (Adelphi, Milano, 1990). Ha curato inoltre la rubrica *Temi metamagici*, pubblicata in «Le Scienze» dal 1981 al 1983.

Bibliografia

Collet Pierre ed Eckmann Jean-Pierre, *Iterated Maps on the Interval as Dynamical Systems*, Birkhauser Boston, Inc., 1980.

Feigenbaum Mitchell J., *Universal Behavior in Nonlinear Systems* in «Los Alamos Science», estate 1980.

7. Alla scoperta delle strane attrattive del caos

Autore

A. K. Dewdney è professore associato di scienza dei calcolatori alla University of Western Ontario. I suoi interessi professionali riguardano soprattutto la matematica del discreto e la teoria della computazione. È autore di *The Planiverse: Computer Contact with a Two-dimensional World*, pubblicato nel 1984 dalla Poseidon Press, e di una raccolta di saggi su vari temi di scienza del calcolatore, *The Turing Omnibus*, edita nel 1985 dalla Computer Science Press. Dal 1984 cura in «Le Scienze» la rubrica *(Ri)creazioni al calcolatore*.

Bibliografia

Hughes Gordon, *Hénon Mapping with Pascal* in «Byte», 11, n. 13, dicembre 1986.

8. Il linguaggio dei frattali

Autori

Hartmut Jürgens, Heinz-Otto Peitgen e Dietmar Saupe collaborano nello studio della matematica di sistemi dinamici complessi, dei frattali e della grafica al calcolatore presso l'Istituto per i sistemi dinamici dell'Università di Brema e hanno fondato un laboratorio di matematica sperimentale in seno all'istituto stesso. Peitgen si è laureato in matematica all'Università di Bonn nel 1973. Ha l'incarico di docente di matematica all'Università di Brema ed è anche *visiting professor* presso

l'Università della California a Santa Cruz. Le sue ricerche si sono concentrate in particolare sull'analisi non lineare e sulle equazioni differenziali, sui metodi numerici, sui sistemi dinamici e sui frattali. Saupe e Jürgens hanno conseguito la laurea in matematica presso l'Università di Brema rispettivamente nel 1982 e 1983. Saupe è assistente di matematica, mentre Jürgens è direttore del laboratorio fondato dal gruppo.

Bibliografia

Mandelbrot B. B., *The Fractal Geometry of Nature*, W. H. Freeman & Co., 1982.

Peitgen H. O. e Richter P., *La bellezza dei frattali*, Bollati Boringhieri, Torino, 1987.

Barnsley M., *Fractals Everywhere*, Academic Press, Inc., 1988.

Peitgen H. O. e Saupe D., *The Science of Fractal Images*, Springer-Verlag, 1988.

Peitgen Heinz-Otto, Jürgens Hartmut e Saupe Dietmar, *Fractals for the Classroom*, Springer-Verlag, 1989.

Peitgen H. O., Jürgens H., Saupe D. e Zahlten C., *Fractals: An Animated Discussion*, con Edward Lorenz e Benoit B. Mandelbrot (video), W. H. Freeman & Co., 1990.

9. L'accrescimento dei frattali

Autore

Leonard M. Sander è professore di fisica all'Università del Michigan. Svolge esperimenti nel campo della teoria dello stato solido e della fisica statistica. Si è diplomato alla Washington University a St. Louis e nel 1969 ha conseguito la laurea in fisica presso l'Università della California a Berkeley. Ha trascorso un anno come borsista all'Università della California a San Diego prima di trasferirsi all'Università del Michigan.

Bibliografia

Witten T. A., Jr., e Sander L. M., *Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon* in «Physical Review Letters», 47, n. 19, 9 novembre 1981.

Mandelbrot Benoit B., *The Fractal Geometry of Nature*, W. H. Freeman and Company, 1982.

Witten T. A. e Sander L. M., *Diffusion-Limited Aggregation* in «Physical Review B», 27, n. 9, 1 maggio 1983.

Witten T. A. e Cates M. E., *Tenuous Structures from Disorderly Growth Processes* in «Science», 232, n. 4758, 27 giugno 1986.

10. Caos e frattali in fisiologia umana

Autori

Ary L. Goldberger, David R. Rigney e Bruce J. West collaborano nello studio della dinamica non lineare in fisiologia umana. Goldberger è assistente di medicina presso la Harvard Medical School e condirettore degli Electrocardiography and Arrhythmia Laboratories al Beth Israel Hospital di Boston. Rigney è assistente di medicina a Harvard e ricercatore al Massachusetts Institute of Technology. West è professore di fisica e preside del Dipartimento di fisica della University of North Texas.

Bibliografia

West B. J., *An Essay on the Importance of Being Nonlinear* in *Lecture Notes in Biomathematics 62*, a cura di S. Levine, Springer-Verlag, 1985.

Goldberger Ary L. e West Bruce J., *Fractals in Physiology and Medicine* in «Yale Journal of Biology and Medicine», 60, pp. 421-435, 1987.

West Bruce J. e Goldberger Ary L., *Physiology in Fractal Dimensions* in «American Scientist», 75, n. 4, luglio-agosto 1987.

Goldberger A. L., Rigney D. R., Mietus J., Antman E. M. e Greenwald S., *Nonlinear Dynamics in Sudden Cardiac Death Syndrome: Heartrate Oscillations and Bifurcations* in «Experientia», 44, pp. 983-987, 1988.

Minetti Alberto E., *I frattali in biomedicina* in *Scienza & Tecnica 89/90*, Arnoldo Mondadori, Milano, 1989.

11. I frattali e le linee suturali delle ammoniti

Autore

Giuseppe Damiani è attualmente ricercatore distaccato presso l'Istituto per la difesa e la valorizzazione del germoplasma animale del CNR a Milano. Nel 1976, 1977 e 1979 è stato uno dei vincitori del Concorso europeo Philips per giovani inventori e ricercatori. Dal 1979 lavora presso il Dipartimento di genetica e microbiologia dell'Università di Pavia dove si occupa di ricerche nel campo della biologia molecolare. Presso questa stessa università si è laureato in scienze biologiche nel 1981 e ha vinto una borsa di studio triennale per la frequenza alla Scuola di perfezionamento in genetica, della quale ha conseguito il diploma nel 1984.

Bibliografia

Mandelbrot Benoit B., *The Fractal Geometry of Nature*, W. H. Freeman, 1982.

Damiani Giuseppe, *Significato funzionale dell'evoluzione dei setti e delle linee di sutura dei nautiloidi e degli ammonoidi* in *Atti del I Convegno Fossili Evoluzione Ambiente*, a cura di Giovanni Pallini, Editore Comitato Centenario Raffaele Piccinini, Pergola, 1986.

Nittmann Johann e Stanley H. Eugene, *Tip Splitting Without Interfacial Tension and Dendritic Growth Patterns Arising from Molecular Anisotropy* in «Nature», 321, n. 6071, 12 giugno 1986.

Hewitt Roger A. e Westermann Gerd E. G., *Function of Complexly Fluted Septa in Ammonoid Shells. I. Mechanical Principle and Functional Models* in «Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen», 172, n. 1, febbraio 1986.

Hewitt Roger A. e Westermann Gerd E. G., *Function of Complexly Fluted Septa in Ammonoid Shells. II. Septal Evolution and Conclusions* in «Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen», 174, n. 2, marzo 1987.

West Bruce J. e Goldberger Ary L., *Physiology in Fractal Dimensions* in «American Scientist», 75, n. 4, luglio-agosto 1987.

12. Monti frattali, piante graftali e grafica al calcolatore

Autore

A. K. Dewdney (*si veda la nota biografica dell'articolo n. 7*).

Bibliografia

Cook Robert L., Porter Thomas e Carpenter Loren, *Distributed Ray Tracing* in «Computer Graphics», 18, n. 3, luglio 1984.
Van de Panne Michiel, *3-D Fractals* in «Creative Computing», 11, n. 7, luglio 1985.

13. Bellezza e profondità: insiemi di Mandelbrot e di Julia

Autore

A. K. Dewdney (*si veda la nota biografica dell'articolo n. 7*).

Bibliografia

Peitgen Heinz-Otto e Richter Peter H., *La bellezza dei frattali*, Bollati Boringhieri, Torino, 1987.

14. La convezione

Autori

Manuel G. Velarde e Christiane Normand collaborano a studi sull'argomento di questo articolo da quando, nel 1974, lavoravano insieme nella divisione di fisica teorica del Centre d'Études Nucléaires di Saclay, in Francia. Ora Velarde è professore di meccanica statistica e direttore del Dipartimento di fisica dei fluidi all'Università autonoma di Madrid. Laureatosi a Madrid nel 1968 e a Bruxelles nel 1970, Velarde ha lavorato a lungo in vari laboratori di ricerca in Francia, Norvegia, Belgio e Inghilterra. La Normand lavora ancora per il Commissariat à l'Énergie Atomique a Saclay, dove si era recata nel 1973 con una borsa di studio subito dopo essersi diplomata in fisica dei plasmi presso l'Università di Parigi-Sud. Dal 1976 è ricercatrice associata del Consiglio nazionale delle ricerche francese (CNRS).

Bibliografia

Chandrasekhar S., *Hydrodynamic and Hydromagnetic Stability*, Oxford University Press, 1961.
Stong C. L. *The Amateur Scientist* in «Scientific American», 216, n. 1, gennaio 1967.
Stong C. L., *The Amateur Scientist* in «Scientific American», 223, n. 3, settembre 1970.
Turner J. S., *Buoyancy Effects in Fluids*, Cambridge University Press, 1973.
Normand Christiane, Pomeau Yves e Velarde Manuel G., *Convective Instability: A Physicist's Approach* in «Review of Modern Physics», 49, n. 3, luglio 1977.
Walker Jearl, *The Amateur Scientist* in «Scientific American», 237, n. 4, ottobre 1977.

15. Come nasce la turbolenza

Autori

Christian Vidal e Jean-Claude Roux lavorano insieme da parecchi anni in un laboratorio del CNRS, il Centre de Recherche Paul Pascal. Professore all'Università di Bordeaux I, Vidal ha iniziato a interessarsi alle reazioni chimiche oscillanti durante un soggiorno presso il laboratorio di Ilya Prigogine a Bruxelles. Roux è assistente all'Università di Bordeaux; dopo aver svolto la tesi sulle reazioni chimiche degli alcoli sottoposti a irradiazione, è entrato a far parte del gruppo di ricerca che studia queste reazioni presso il Centre de Recherche Paul Pascal. In questo campo, il contributo del gruppo degli autori è consistito soprattutto nell'identificazione e nella misurazione di specie chimiche intermedie, oltre che nella determinazione degli effetti dell'agitazione e del trasporto di calore. Le apparecchiature e i metodi messi a punto in questo lavoro hanno consentito di ottenere per la prima volta risultati sperimentali quantitativi sulla turbolenza chimica.

Bibliografia

Winfree Arthur T., *Reazioni chimiche rotanti* in «Le Scienze» n. 74, ottobre 1974.
Tysson John J., *The Belousov-Zhabotinsky Reaction*, Springer-Verlag, Berlino-Heidelberg, 1976.
Swinney Harry L. e Gollub Jerry P., *The Transition to Turbulence* in «Physics Today», 31, p. 41, agosto 1978.

16. Il mescolamento dei fluidi

Autore

Julio M. Ottino è professore di ingegneria chimica e professore aggregato di scienza e tecnologia dei polimeri presso l'Università del Massachusetts ad Amherst. Ha compiuto gli studi all'Università nazionale di La Plata, in Argentina, e all'Università del Minnesota, dove si è laureato nel 1979. Le sue esperienze nel mescolamento di materiali colorati non si limitano al laboratorio; è infatti appassionato di pittura e ha recentemente esposto i suoi quadri in una mostra personale in Argentina. Ottino ha recentemente completato il libro *The Kinematics of Mixing: Stretching, Chaos, and Transport* che verrà pubblicato dalla Cambridge University Press.

Bibliografia

Aref H., *Stirring by Chaotic Advection* in «Journal of Fluid Mechanics», 143, giugno 1984.
Ranz William E., *Fluid Mechanical Mixing-Lamellar Description* in *Mixing of Liquids by Mechanical Agitation*, a cura di Jaromir J. Ulbrecht e Gary K. Patterson, Gordon and Breach Science Publishers, 1985.
Khakhar D. V., Rising H. e Ottino J. M., *Analysis of Chaotic Mixing in Two Model Systems* in «Journal of Fluid Mechanics», 172, pp. 419-451, novembre 1986.
Ottino J. M., Leong C. W., Rising H. e Swanson P. D., *Morphological Structures Produced by Mixing in Chaotic Flows* in «Nature», 333, n. 6172, 2 giugno 1988.

17. I vetri di spin e lo studio dei mezzi disordinati

Autori

J. Hammann e M. Ocio hanno iniziato la loro carriera scientifica rispettivamente presso il Servizio di fisica dei solidi del Centre d'Études Nucléaires di Saclay nel 1964 e presso il Commissariat à l'Énergie Atomique nel 1966. Hammann ha studiato fra l'altro le strutture magnetiche indotte a bassissima temperatura, in taluni solidi, dalle interazioni di dipolo utilizzando tecniche di diffrazione neutronica. È anche animatore di un

gruppo di ricerca sulle proprietà magnetiche della materia condensata. Il crescente interesse suscitato dallo studio dei mezzi disordinati lo ha condotto a occuparsi da vicino, a partire dal 1981, delle proprietà dei vetri di spin. Ocio ha effettuato i suoi primi lavori nel campo dei semiconduttori a piccola banda proibita. Nel 1974 è entrato a far parte del Servizio di fisica dei solidi dell'Institut de Recherche Fondamentale del CEA, dove si occupa delle proprietà magnetiche della materia condensata a temperature ultrabasse. La sua esperienza nel campo della magnetometria fine lo ha condotto, in collaborazione con H. Bouchiat e P. Monod dell'Università di Orsay, a mettere a punto il primo esperimento mai compiuto di rilevamento del rumore magnetico spontaneo emesso da composti magnetici.

Bibliografia

Rammal R. e Souletie J., *Magnetism of Metals and Alloys*, Cinor, Amsterdam, 1982.

Mézard M., *La physique statistique des verres de spin*, Images de la physique, 1985.

Ocio M., Bouchiat H. e Monod P., *Le bruit magnétique des verres de spin*, Aspects de la recherche, Université Paris-Sud, 1986.

18. «Vetri di spin»

Autore

Daniel L. Stein è professore associato di fisica all'Università dell'Arizona. Ha compiuto gli studi superiori alla Brown University e successivamente alla Princeton University, presso la quale si è laureato nel 1979. È stato il primo direttore della Summer School on Complex Systems di Santa Fe (New Mexico) nel 1988. Gli argomenti di cui si è occupato nelle sue ricerche comprendono la complessità nel calcolo, i vetri di spin, i vetri strutturali e la biofisica.

Bibliografia

Mézard M., Parisi G. e Virasoro M. A., *Spin Glass Theory and Beyond*, World Scientific Publications, Singapore, 1986.

Chowdhury D., *Spin Glasses and Other Frustrated Systems*, World Scientific Publications, Singapore, 1986.

Stein D. L. (a cura), *Complex Systems*, Addison-Wesley Publishing Company, 1989.

Kauffman S. A., *Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, 1989.

19. Reazioni chimiche oscillanti

Autori

Irving R. Epstein, Kenneth Kustin, Patrick De Kepper e Miklós Orbán sono chimici. Epstein è docente di chimica alla Brandeis University. Si è diplomato allo Harvard College e laureato alla Harvard University; è alla Brandeis dal 1971. Nel 1977 e 1978 è stato borsista della National Science Foundation al Max Planck Institut für Biophysikalische Chemie di Göttingen. Anche Kustin, che si è diplomato al Queens College nel 1955 e si è laureato in chimica inorganica all'Università del Minnesota, è docente di chimica alla Brandeis University. Dopo essere stato borsista dell'US Public Health Service presso il Max Planck Institut für Physikalische Chemie di Berlino, si è trasferito alla Brandeis University nel 1961, e dal 1974 al 1977 è stato direttore del Dipartimento di chimica. De Kepper è direttore delle ricerche in ingegneria chimica al Centre de Recherche Paul Pascal del CNRS. Ha studiato all'Università di Bordeaux I, laureandosi in ingegneria chimica nel 1978, per passare poi, nello stesso anno, al CNRS. Orbán è nato in Ungheria e si è laureato all'Università Eötvös di Budapest, dove insegna dal 1962. Attualmente è professore associato di chimica.

Bibliografia

Field Richard J., Körös Endre e Noyes Richard M., *Oscillations in Chemical Systems, II: Thorough Analysis of Temporal Oscillation in the Bromate-Cerium-Malonic Acid System* in «Journal of the American Chemical Society», 94, n. 25, 13 dicembre 1972.

Chance B., Pye E. K., Ghosh A. K. e Hess B. (a cura), *Biological and Biochemical Oscillators: Proceedings of the Conference on Biological and Biochemical Oscillators*, Academic Press, New York, 1973.

Nicolis G. e Portnow J., *Chemical Oscillations* in «Chemical Reviews», 73, n. 4, agosto 1973.

De Kepper Patrick, Epstein Irving R. e Kustin Kenneth, *A Systematically Designed Homogeneous Oscillating Reaction: The Arsenite-Iodate-Chlorite System* in «Journal of the American Chemical Society», 103, n. 8, 22 aprile 1981.

Orbán Miklós, De Kepper Patrick ed Epstein Irving R., *Minimal Bromate Oscillator: Bromate-Bromide-Catalyst* in «Journal of the American Chemical Society», 104, n. 9, 5 maggio 1982.

20. Caos quantistico?

Autore

Questo pezzo è stato scritto dalla redazione di «Scientific American».

21. La casualità in aritmetica

Autore

Gregory J. Chaitin fa parte del Thomas J. Watson Research Center dell'IBM a Yorktown Heights, New York. È il principale ideatore della teoria algoritmica dell'informazione e ha pubblicato due libri nei quali sono applicati i principi di questa teoria per chiarire la natura della casualità e le limitazioni della matematica. È autore dell'articolo *Casualità e dimostrazione matematica*, pubblicato in «Le Scienze» n. 85, settembre 1975.

Bibliografia

Chaitin Gregory J., *Algorithmic Information Theory*, Cambridge University Press, 1987.

Chaitin Gregory J., *Information, Randomness & Incompleteness*, World Scientific Publishing Co. Pte. Ltd., 1987.

Stewart Ian, *The Ultimate in Undecidability* in «Nature», 232, n. 6160, 10 marzo 1988.

22. Diavoletti, macchine e il secondo principio

Autore

Charles H. Bennett fa parte del gruppo di ricercatori operanti presso il Thomas J. Watson Research Center della IBM

Corporation a Yorktown Heights. Si è laureato in chimica fisica presso la Harvard University nel 1970 e ha poi lavorato per due anni all'Argonne National Laboratory: si è trasferito all'IBM nel 1972. Nelle sue ricerche si è interessato alla teoria matematica della casualità e alle relazioni fra meccanica statistica e teoria del calcolo. Bennett è coautore dell'articolo *I limiti fisici fondamentali del calcolo*, pubblicato in «Le Scienze» n. 205, settembre 1985.

Bibliografia

Landauer R., *Irreversibility and Heat Generation in the Computing Process* in «IBM Journal of Research and Development», 5, n. 3, luglio 1961.

Brillouin Léon, *Science and Information Theory*, Academic Press, 1962.

Bennett Charles H., *The Thermodynamics of Computation--a Review* in «International Journal of Theoretical Physics», 21, n. 12, dicembre 1982.

Szilard Leo, *On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings* in *Quantum Theory and Measurement*, a cura di John Archibald Wheeler e Wojciech Hubert Zurek, Princeton University Press, 1983.

Finito di stampare nel mese di dicembre 1990
dalla Milanostampa, Farigliano (CN).

L. 26.000